e-Business and Telecommunication Networks

Edited by
João Ascenso, Luminita Vasiu, Carlos Belo
and Mónica Saramago

Springer

e-Business and Telecommunication Networks

e-Business and Telecommunication Networks

edited by

João Ascenso

ISEL,
Lisbon, Portugal

Luminita Vasiu

University of Westminster,
London, UK

Carlos Belo

IST/IT,
Lisbon, Portugal

and

Mónica Saramago

INSTICC,
Setúbal, Portugal

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-4760-6 (HB)
ISBN-13 978-1-4020-4760-2 (HB)
ISBN-10 1-4020-4761-4 (e-book)
ISBN-13 978-1-4020-4761-9 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

Printed in the Netherlands

TABLE OF CONTENTS

PART 1 – GLOBAL COMMUNICATION INFORMATION SYSTEMS AND SERVICES

PART 2 – SECURITY AND RELIABILITY IN INFORMATION SYSTEMS AND NETWORKS

PREFACE

This book contains the best papers of the First International Conference on E-business and Telecommunication Networks (ICETE 2004), held in Setúbal (Portugal) and organized by INSTICC (*Institute for Systems and Technologies of Information, Communication and Control*) in collaboration with the School of Business of the Polytechnic Institute of Setúbal, who hosted the event.

This conference represents a major initiative to increase the technical exchanges among professionals, who work on the e-Business and Telecommunication Networks fields, and who are deploying new services and technologies into the lives of ordinary consumers. The major goal of this conference is to bring together researchers and developers from academia and industry working in areas related to e-business, with a special focus on Telecommunication Networks. This year, four simultaneous tracks were held, covering different aspects, including: "*Global Communication Information Systems and Services*", "*Security and Reliability in Information Systems and Networks*", "*Wireless Communication Systems and Networks*" and "*Multimedia Signal Processing*". The sections of this book reflect the conference tracks.

ICETE 2004 received 202 paper submissions from 43 different countries, from all continents. 110 papers were published and orally presented as full papers, i.e. completed work, and 44 papers were accepted for poster presentation. The full paper acceptance ratio confirms our confidence that ICETE 2004 has achieved a high quality standard that we will strive to keep and enhance in order to ensure the success of the next year ICETE edition.

Additionally, the ICETE conference included a number of invited talks, including keynote lectures and technical tutorials. These special presentations made by internationally recognized experts have definitely increased the overall quality of the Conference and provided a deeper understanding of the Telecommunication Networks field. Their contributions have been included in a special section of this book.

The program for this conference required the dedicated effort of many people. Firstly, we must thank the authors, whose research and development efforts are recorded here. Secondly, we thank the members of the program committee and the additional reviewers for their diligence and expert reviewing. Thirdly, we thank the invited speakers for their invaluable contribution and for taking the time to synthesise and prepare their talks. Finally, we thank the workshop chairs whose collaboration with ICETE was much appreciated.

João Ascenso
School of Technology of Setúbal, IPS, Setúbal, Portugal

Luminita Vasiu
Middlesex University, WITRC, London, U.K.

Joaquim Filipe
School of Technology of Setúbal and INSTICC, Setúbal, Portugal

Carlos Belo
Institute of Telecommunications and IST, Lisbon, Portugal

CONFERENCE COMMITTEE

Conference Chair
Joaquim Filipe, Escola Superior de Tecnologia de Setúbal, Portugal

Programme co-Chairs
Carlos Belo, Instituto de Telecomunicações, Portugal
Luminita Vasiu, Middlesex University, U.K.

Program Committee Chair
João Ascenso, Escola Superior de Tecnologia de Setúbal, Portugal

Secretariat
Mónica Saramago, INSTICC, Portugal

Programme Committee:
Acharya, A. (USA)
Ahmed, K. (THAILAND)
Al-Sharhan, S. (KUWAIT)
Ansari, N. (USA)
Asatani, K. (JAPAN)
Assuncão, P. (PORTUGAL)
Barn, B. (UK)
Bedford, A. (AUSTRALIA)
Bella, G. (ITALY)
Benzekri, A. (FRANCE)
Boavida, F. (PORTUGAL)
Bonyuet, D. (USA)
Boutaba, R. (CANADA)
Broadfoot, P. (UK)
Cappellini, V. (ITALY)
Cheng, T. (SINGAPORE)
Cheung, K. (CHINA)
Choras, R. (POLAND)
Clarke, R. (UK)
Cohen, R. (ISRAEL)
Comley, R. (UK)
Constantinides, T. (UK)
Correia, M. (PORTUGAL)
Correia, P. (PORTUGAL)
Devetsikiotis, M. (USA)
Elmirghani, J. (UK)
Fang, Y. (USA)
Faria, S. (PORTUGAL)
Figueiredo, M. (PORTUGAL)
Gaspary, L. (BRAZIL)
Georghiades, C. (USA)
Giannakis, G. (GREECE)
Goldszmidt, G. (USA)
Goulart, C. (BRAZIL)
Granai, L. (SWITZERLAND)
Granville, L. (BRAZIL)
Greaves, D. (UK)
Gritzalis, S. (GREECE)
Kang, C. (KOREA)
Hamdi, M. (CHINA)
Hanzo, L. (UK)
Harris, R. (AUSTRALIA)
Helal, S. (USA)
Hoang, N. (SINGAPORE)
Hong, D. (KOREA)
Hu, J. (AUSTRALIA)
Huston, G. (AUSTRALIA)
Isaías, P. (PORTUGAL)
Jagodic, M. (SLOVENIA)
Jahankhani, H. (UK)
Jain, A. (INDIA)
Jefferies, N. (UK)
Júnior, E. (BRAZIL)
Kahlil, I. (AUSTRALIA)

Karmouch, A. (CANADA)
Kihl, M. (SWEDEN)
Kollias, S. (GREECE)
Kos, M. (CROATIA)
Kunt, M. (SWITZERLAND)
Kuo, G. S. (TAIWAN)
Landfeldt, B. (AUSTRALIA)
Lee, M. (AUSTRIA)
Lewis, L. (USA)
Liu, K. (UK)
Lloyd-Smith, B. (AUSTRALIA)
Lorna, U. (UK)
Loureiro, A. (BRAZIL)
Lu, S. (USA)
Magedanz, T. (GERMANY)
Magli, E. (ITALY)
Mahmoud, Q. (CANADA)
Makki, K. (USA)
Malek, M. (USA)
Malumbres, M. (SPAIN)
Man, H. (USA)
Marshall, A. (UK)
Marshall, I. (UK)
Mascolo, S. (ITALY)
Matsuura, K. (JAPAN)
McGrath, S. (IRELAND)
Merabti, M. (UK)
Mirmehdi, M. (UK)
Morikawa, H. (JAPAN)
Navarro, A. (PORTUGAL)
Nordholm, S. (AUSTRALIA)
Obaidat, M. (USA)
Ohtsuki, T. (JAPAN)
Osadciw, L. (EUA)
Pach, A. (POLAND)
Perkis, A. (NORWAY)
Petrizzelli, M. (VENEZUELA)
Pigneur, Y. (SWITZERLAND)
Pinnes, E. (USA)
Pitsillides, A. (CYPRUS)
Plagemann, T. (NORWAY)
Podvalny, S. (RUSSIA)
Preston, D. (UK)
Queluz, P. (PORTUGAL)
Ramadass, S. (MALAYSIA)
Raychaudhuri, D. (USA)
Regazzoni, C. (ITALY)
Reichl, P. (AUSTRIA)
Reis, L. (PORTUGAL)
Rodrigues, A. (PORTUGAL)
Rosales, C. (MEXICO)
Roth, J. (GERMANY)
Roztocki, N. (USA)
Sanadidi, M. (USA)
Schulze, B. (BRAZIL)
Sericola, B. (FRANCE)
Skarbek, W. (POLAND)
Specialski, E. (BRAZIL)
Steinmetz, R. (GERMANY)
Suda, T. (USA)
Sun, L. (UK)
Sure, Y. (GERMANY)
Tarouco, L. (BRAZIL)
Tirri, H. (FINLAND)
Toh, C. K. (USA)
Ultes-Nitsche, U. (SWITZERLAND)
Valadas, R. (PORTUGAL)
Vidal, A. (SPAIN)
Waldron, J. (IRELAND)
Weghorn, H. (GERMANY)
Weigel, R. (GERMANY)
Wilde, E. (SWITZERLAND)
Wilson, S. (USA)
Wu, G. (USA)
Wu, W. X. (UK)
Yasinsac, A. (EUA)
Yeo, B. (SINGAPORE)
Yin, Q. (SINGAPORE)
Youn, H. (KOREA)
Yu, W. (USA)
Yuan, S. (TAIWAN)
Zhang, J. (USA)

Invited Speakers

Luminita Vasiu, Middlesex University, U.K.
Manu Malek, Institute if Technology, USA
Henry Tirri, Nokia Research Fellow/Nokia Research Center, Finland
Nirwan Ansari, New Jersey Institute of Technology, USA
Mohamed Atiquzzam, Uiversity of Oklahoma, USA

Invited Speakers

DATA MINING TECHNIQUES FOR SECURITY OF WEB SERVICES

Manu Malek and Fotios Harmantzis
Steven Institute of Technology, Castle Point on the Hudson, Hoboken, NJ 07030, USA
Email: {mmalek, fharmant}@stevens.edu

Keywords: Security services, Security attack, Denial of service, Intrusion detection, Security safeguards.

Abstract: The Internet, while being increasingly used to provide services efficiently, poses a unique set of security issues due to its openness and ubiquity. We highlight the importance of security in web services and describe how data mining techniques can offer help. The anatomy of a specific security attack is described. We then survey some security intrusions detection techniques based on data mining and point out their shortcomings. Then we provide some novel data mining techniques to detect such attacks, and describe some safeguard against these attacks.

1 INTRODUCTION

Cyberspace is used extensively for commerce. For years banks and other financial organizations have conducted transactions over the Internet using various geographically dispersed computer systems. Businesses that accept transactions via the Internet can gain a competitive edge by reaching a worldwide customer base at relatively low cost. But the Internet poses a unique set of security issues due to its openness and ubiquity. Indeed, security is recognized as a critical issue in Information Technology today. Customers will submit information via the Web/ Internet only if they are confident that their private information, such as credit card numbers, is secure. Therefore, today's Web/Internet-based services must include solutions that provide security as a primary component in their design and deployment.

Web services generally refer to web-based applications that make it possible for enterprises to do transactions on the web and for users to share documents and information with each other over the Web. The standard that makes it possible to describe the communications in some structured way is Web Services Definition Language (WSDL). WSDL is an XML format for describing network services as a set of endpoints operating on messages containing either document-oriented or procedure-oriented information (http://www.w3.org/TR/wsdl).

But openness and integration have their price. Without adequate security protections and effective security management, these features can be used to attack the availability and integrity of information systems and the networks connecting to them. Here we highlight a few typical ways an attacker may gain illegal access to an information system, or to make it unavailable to legitimate users. We identify the profiles or signatures for the sequence of actions an attacker may perform to perpetrate such attacks. We use data mining techniques to discover such attack profiles to detect the attacks.

Data mining refers to a technique to intelligently and automatically assist humans in analyzing the large volumes of data to identify valid, novel, and potentially useful patterns in data. It offers great promise in helping organizations uncover patterns hidden in their data that can be used to predict the behavior of customers, so that they can better plan products and processes. Data mining takes advantage of advances in the fields of artificial intelligence (AI) and statistics. Both disciplines help in pattern recognition and classification. Other disciplines used in data mining include rule-based and case-based reasoning, fuzzy logic, and neural networks. The techniques used in data mining include rule induction, clustering, projection, and visualization (e.g., see (Berry, M. and L. Gordon, 1997) for details).

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 3–10.

This paper provides a glimpse at the cyberspace security situation, and offers some techniques to manage the security of web services. The paper describes some security attacks, and provides some techniques to detect and defend against them. In Section 2, we present some statistics related to security attacks to highlight the urgency of the issue. Some typical security vulnerabilities and attacks are discussed in Section 3. In Section 4, we provide a survey of data mining applications in intrusion detection and point out their shortcomings. We then define attack signatures and outline how to use them in conjunction with data mining techniques for efficient intrusion detection. Section 5 summarizes the paper.

2 BACKGROUND

Based on data provided by CERT/CC, the number of incidents and vulnerabilities for cyber attacks have increased exponentially during the period 1998 to 2002 (CERT/CC). Figure 1 shows that intrusions were relatively few in the early1990s, but there has been a major increase since 2000. About 25,000 intrusions were reported in the Year 2000 (CERT/CC). Keep in mind that not all enterprises that suffer security breaches report them. The line moving upward in this figure shows various types of threats, starting with very simple ones in the early '90s, like password guessing. The sophistication of attacks increased with self-replicating codes, such as viruses, then password cracking (where the cryptographic password is broken), and on to the more sophisticated threats shown. Against this rising sophistication in threats, we have easy availability of hacking tools: hackers no longer have to be experts in computer science or security; they could use available tools. For example, a tool such as nmap (www.insecure.org/nmap/nmap-fingerprinting-aticle.html) can be used to find all the open ports, a first stage in an attack. This combination of decreasing knowledge required of the attackers and the increasing sophistication of the attacks is giving rise to major security concerns.

According to the Federal Computer Incidence Response Center (FedCIRC), the incident handling entity for the federal government, 130,000 government sites totaling more than one million hosts were attacked in 1998 (NIST ITL Bulletin, 1999). Also, a 1999 survey conducted by the Computer Security Institute (CSI) and the Federal Bureau of Investigation (FBI) revealed that 57% of organizations cited their Internet connections as a frequent point of attack, 30% detected actual network intrusion, and 26% reported theft of proprietary information (CSI, 2002). A similar survey in 2002, showed that 90% of the 507 participating organizations detected computer security breaches within the past 12 months, 74% cited their Internet connections as a frequent point of attack, but only 34% reported security intrusions to law enforcement agencies. These numbers must be considered observing that they relate to only known attacks and vulnerabilities. However, they do indicate the magnitude of the problem.

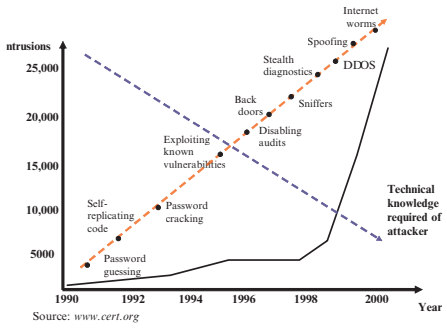


Figure 1: Security vulnerabilities and threats.

A key to preventing security attacks is to understand and identify vulnerabilities, and to take corrective action. A threat to computing systems or communication network is a potential violation of security unauthorized, illegitimate, malicious or fraudulent purposes. An attack is the implementation of a threat using the system vulnerabilities. Vulnerability is a weakness in the security system that might be exploited to launch an attack. Finally, a control is a protective measure – an action, device, procedure, or technique – that reduces vulnerabilities.

Table 1 sows the top 10 security vulnerabilities as reported periodically by The SANS Institute (The SANS Institute, 2003). The reasons for the existence of these vulnerabilities include: buggy software design and development, system administrators being too busy to install security patches in a timely manner, and inadequate policies and procedures. Another factor is that due to the ubiquity of the Internet, vulnerabilities are quickly and widely published.

Table 1: Top 10 security vulnerabilities (The SANS Institute, 2003).

Vulnerabilities of Windows Systems	***Vulnerabilities of Unix Systems***
1. Internet Information Services (IIS)	*1. Remote Procedure Calls (RPC)*
2. Microsoft Data Access Components (MDAC)– Remote Data Services	*2. Apache Web Server*
3. Microsoft SQL Server	*3. Secure Shell (SSH)*
4. NETBIOS – Unprotected Windows Networking Shares	*4. Simple Network Management Protocol (SNMP)*
5. Anonymous Logon – Null Sessions	*5. File Transfer Protocol (FTP)*
6. LAN Manager Authentication – Weak LM Hashing	*6. R-Services – Trust Relationships*
7. General Windows Authentication – Accounts with No Passwords or Weak Passwords	*7. Line Printer Daemon (LPD)*
8. Internet Explorer	*8. Sendmail*
9. Remote Registry Access	*9. BIND/DNS*
10. Windows Scripting Host	*10. General Unix Authentication – Accounts with No Passwords or Weak Passwords*

The motive of attackers could be anything from pure joy of hacking to financial benefit. The attackers are either highly technically capable, or they sometimes breaks into the network by trial and error. Disgruntled employees have more access rights to enterprise computer networks compared to outside attackers. According to the CSI/FBI 2002 Survey (CSI, 2002), 60% of attacks in the US were inside attacks (attacks that originated inside the institutions) and 40% were outside attacks.

3 SOME TYPICAL SECURITY ATTACKS

As mentioned, a security attack occurs when an attacker takes advantage of one or more security vulnerabilities. To improve security, one needs to minimize security vulnerabilities. In this section we present some typical security attacks, point out the vulnerabilities abused to perpetrate the attacks. Some safeguards against these attacks will be described in the next section. An attack that we deal with specifically in this section and in Section 4 is the HTTP GET attack.

3.1 Denial-of-Service Attack

In a Denial-of-Service (DoS) attack, the attacker attempts to use up all the victim system's resources like memory or bandwidth. When the attack is successful, legitimate users can no longer access the resources and the services offered by the server will be shut down. According to the 2002 CSI/FBI survey (CSI, 2002), 40% of all attacks are DoS attacks.

An attack can be directed at an operating system or at the network. The attacker may send specially crafted packets that crash remote software/services running on the victim server. It will be successful if the network is unable to distinguish between legitimate traffic and malicious or bogus traffic. Some common DoS attacks follow.

ICMP Flooding and Smurf Attack
These are both ICMP-based attacks. Flooding with ICMP packets slows down the victim server so that it can no longer respond quickly enough for the services to work properly. If packets are sent with forged IP addresses, the victim server not only has to allocate system resources to receive, but to reply to packets to addresses which do not exist. The Smurf attack uses a similar idea: the attacking machine sends Echo requests with broadcast IP addresses, thus not only the victim server but the attached network will be flooded by a large amount of ICMP traffic.

SYN Flooding
SYN flooding exploits the weakness of the TCP Three-way Handshake (Comer, D., 2000). In a normal TCP connection request, the source sends a SYN (synchronization) packet to the destination to initiate the connection; then waits for a SYN ACK (synchronization acknowledged) packet from the destination. The connection is established when the destination receives a FIN ACK (finishing acknowledged) packet from the source. In the SYN flooding attack, the attacker sends a large number of SYN packets, often from bogus IP addresses, to the victim server, which adds the entry to the connection queue and replies with SYN ACKs. As the source addresses are incorrect or non-existent, FIN ACKs

will never be received by the victim server, so the last part of the Three-way Handshake never completes and the connection queue of the victim server fills up.

Badly-formed Packets
In this type of attack, the attacker sends badly-formed IP packets, e.g., packets that consist of invalid fragments, protocol, packet size, or header values, to the victim server. Once the destination TCP stack receives such invalid packets, the operating system must allocate resources to handle them. If the operating system cannot handle the misbehavior, it will crash. An example of this is the Ping-of-Death attack (www.insecure.org/sploits/ping-o-death.html) which causes buffer overflow in the operating system. In this attack, the attacker sends a larger than standard ICMP (Internet Control Message Protocol) packet, such as a ping, in fragments to the target server. Since the allowed maximum size of such a packet is 65,535 bytes, the server allows a corresponding buffer space to collect the fragments. A clever attacker may create a ping with many fragments destined to a target server. The server receives the fragments and starts to reassemble them. When reassembled, the buffer will overflow, leading to program termination, overwriting other data or executable code, kernel dump, etc. More than 50% of attacks on servers are due to buffer overflow (CERT/CC).

Distributed Denial of Service Attack
With the speed and power of computing resources today, an attacker may not be able to simply use one computer to craft a DoS attack. In the Distributed Denial of Service (DDoS) attack, many computers may be hijacked by the attacker as agents (zombies) to simultaneously flood a victim system's resources. A typical way to recruit zombie computers is for the perpetrator to send viruses to multiple computers, or to break into computer systems and load them with DDoS programs. Each infected system then finds other vulnerable systems and loads them with the programs, etc. The perpetrator uses the first system that was overtaken to instruct all the other compromised systems to launch the attack simultaneously.

3.2 HTTP GET Attack

For many web applications, a client should be able to send information to the server. HTML 2.0 and later versions support the Form element within an HTML document to allow data to be sent to web servers (www.w3c.org). One of the attributes of Form is Method which indicates how data is submitted to the web server. Valid choices for the Method attribute are GET and POST. In METHOD = GET the values inputted by the user are concatenated with the URL, separated by a special character (usually ?) fields are separated by &; space is represented by + For example, the following URL: http://www.gadgets, com?customer = **John + Doe** & address = **101 + Main + Street** & cardno = **1234567890** & cc=; **visacard** indicates that the customer's name and address with the customer's credit card number are to be sent to the web server at www.gadgets.com A savvy user (attacker) may be able to use this feature to get access to proprietary information if appropriate security mechanisms are not in place. The following scenario, adopted from Ref. (McClure, S. et al., 2003), is an HTTP GET attack on a typical web server which has some vulnerabilities.

The server http://www.acme.com runs Apache 1.3.12 on a Linux operating system. Firewalls prevent all but HTTP traffic via ports 80 (HTTP default port) and 443 (SSL port). Perl CGI scripts are used for the online store. A visitor to this site first begins browsing through the www.acme.com site, viewing the site's main page and a few images on it. The visitor notices that for the last selection (viewing the picture of a sunset), the URL in the browser window shows: http://www.acme.com/index.cgi?page=sunset.html. Following this pattern, the visitor (now attacker) issues a request for index.cg by typing the following URL: http://www.acme.com/index.cgi?page=index.cgi

Now, if the program does not validate the parameters passed to the index.cgi script, the filename passed as a parameter from the URL is captured by the CGI script, appended to the absolute path, and causes to open the index.cgi script as requested. Consequently, the browser display shows the source code of the index.cgi script!

At this point the attacker realizes that this technique can be further exploited to retrieve arbitrary files from the server. So the attacker may send the following request through the browser: http://www.acme.com/index.cgi?page=/../../etc/passwd

If permissions are not set properly on the /etc/passwd file, its contents will be displayed by the browser, providing the attacker with user's password information. The attacker could now execute arbitrary commands on the server, for example, by sending http://www.acme.com/index.cgi?page=|ls+-la/%0aid%0awhich+xterm| (% plus the hex character

0a indicates line feed) requesting ls –al (to show a file list of the server's root directory) id (the effective user id of the process running index.cgi) which xterm (path to the xterm binary code, to gain interactive shell access to the server and the attacker could gain full interactive shell-level access to the web server.

Note again that the vulnerabilities: the program does not validate the parameters passed to the index.cgi script, and permissions are not set properly on the /etc/passwd file.

4 DATA MINING TECHNIQUES FOR INTRUSION DETECTION

In this section we review server logs, introduce attack signatures, and present our main contribution: how security attack signatures are used in conjunction with data mining to detect security intrusions. More specifically, we first describe the relevance and importance of the different log files that are available; we then define specific patterns in the log files for an attack (the individual log records as well as their sequence/order) as the attack signature; and use data mining to search and find such patters for attack detection. The efficiency and speed of the overall process can even lead to attack prediction capabilities.

4.1 Logs

Every visit to a Web site by a user creates a record of what happens during that session in the server's log. A busy site may generate thousands of log entries per hour, compiled in various log files. A log file entry contains items like the IP address of the computer requesting the Web page, the date and time of the request, the name and the size of the file requested. Logs vary by the type of server and the file format. Following are some typical logs and what they record:

- Access Log records every transaction between server and browsers (date, time, domain name or IP address, size of transaction, …).
- Referrer Log records the visitor's path to the site (the initial URL from which the visitor came).
- Agent Log records the type and version of the browser.

For secure systems, the standard logs and directories may not be sufficient and one must employ additional logging tools, e.g., information about which computer is connecting to which services on the system. There are many programs under the heading of IP loggers available for this purpose, e.g., EnviroMon (http://www.interwld.com/pico/subs/pico_Environ_IP_Logging.htm) and ippl http://packages.debian.org/unstable/net/ippl.html).

4.2 Mining Logs

The data available in log files can be "mined" to gain useful information. Data mining offers promise in uncovering hidden patterns in the data that can be used to predict the behavior of (malicious) users. Using data mining in intrusion detection is a relatively new concept. In (Lee, W. and Stolfo, S. J., 1998), the authors outline a data mining framework for constructing intrusion detection models. The central idea is to utilize auditing programs to extract an extensive set of features that describe each network connection or host session, and apply data mining to learn rules that capture the behavior of intrusions and normal activities. Detection models for new intrusions are incorporated into an Intrusion Detection Systems (IDS) through a meta-learning (or co-operative) learning process. The strength of this approach is in classification, meta-learning, and association rules.

In (Almgren, M. et al., 2000), the authors present an intrusion detection tool aimed at protecting servers. However, their method does not effectively handle all matches of the signature (e.g., of the 404 type: document not found). Attacks that have no matching signature and are sent by a previously unknown host may be missed.

Ref. (Forrest, S. et al., 1996 represents a first attempt to analyze sequences of system calls issued by a process for intrusion detection. The authors introduce a method based on sequences of Unix system calls at the process level for anomaly detection resulting in intrusion detection. They address sendmail, lpr, and ftpd processes and obtain some good results in terms of false-positives. Ref. (Hofmeyr, S.A., 1998) also uses sequences of system calls for intrusion detection. The authors choose to monitor behaviour at the level of privileged processes. Their proposed approach of detecting irregularities in the behavior of privileged programs is to regard the program as a black-box, which, when it runs, emits some observable behavior. Privileged processes

are trusted to access only relevant system resources, but in cases where there is a programming error in the code that the privileged process is running, or if the privileged process is incorrectly configured, an ordinary user may be able to gain super-user privileges by exploiting the problem in the process. The system-call method, however, is specific to processes and cannot detect generic intrusion attempts, e.g., race condition attacks, session hijackin g(when one user masquerades as another), and cases in which a user violates policy without using privileged processes.

Artificial intelligence techniques have also been applied to help in decision making for intrusion detection. In (Frank, J., 1994), the author presents a survey of such methods and provides an example of using feature selection to improve the classification of network connections. In (Liu, Z. et al., 2002), the authors present a comparison of some neural-network-based method and offer some "classifiers" for anomaly detection in Unix processes. All the techniques based on artificial intelligence, however, suffer from lack of scalability: they work only for small size networks and data sizes.

4.3 Attack Signatures

We use attack signatures in combination with data mining to not only detect, but predict attacks. An attack signature encapsulates the way an attacker would navigate through the resources and the actions the attacker would take. For example, in a denial-of-service attack, the attacker may send a large number of almost simultaneous TCP connect requests from one or more IP addresses without responding to server acknowledgements.

To illustrate a specific attack signature, let us look at the log lines stored by the web server in the HTTP GET attack example described in the previous section. The log line in the Access Log corresponding to the visitor's (attacker's) first attempt is

A. 10.0.1.21 – [31/Oct/2001:03:02:47] "GET/HTTP /1.0" 200 3008 where 10.0.1.21 is the visitor's IP address, followed by date and time of visit, the Method and the Protocol used. The number 200 indicates the "normal" code, and 3008 indicates the byte size of the file retrieved. The following log line corresponds to the visitor's selection of the sunset picture:

B. 10.0.1.21 – [31/Oct/2001:03:03:18] "GET/sunset. jpg HTTP/1.0" 200 36580 and the following log line corresponds to the visitor's first attempt at surveillance of the site (issuing a request for index.cgi):

C. 10.0.1.21 – [31/Oct/2001:03:05:31] "GET/index. cgi?page= index.cgi HTTP/1.0" 200 358 The following log lines correspond to the visitor (by now, attacker) attempting to open supposedly secure files:

D. 10.0.1.21 – [31/Oct/2001:03:06:21] "GET/index. cgi?page=/../../etc/passwd HTTP/1.0" 200 723

E. 10.0.1.21 – [31/Oct/2001:03:07:01] "GET/ index. cgi?page=|ls+-la+/%0aid%0awhich+xterm|HTTP/ 1.0" 200 1228

This pattern of log lines from the same source IP address can be recognized as a signature of an HTTP GET attack. In the above example, the sequence of log lines A-B-C-D-E, A-C-D-E, B-C-D-E, or C-D-E constitutes the signature of this HTTP GET attack. Even some individual log lines from a source IP address could provide tell-tale signs of an impending HTTP GET attack. For example, the existence of a "pipe" (i.e.,) in the URL, as in log line E above, would indicate that the user is possibly trying to execute operating system commands.

In our research, we try to establish signatures for various types of attacks. Note the importance of good comprehensive attack signatures in detecting attacks. Incomplete signatures result in false-positive or false-negative detection. Another point is the use of data mining to detect attack signatures. One can imagine the tremendous amount of data collected by web services, resulting in multi-tera-byte databases. With such large amounts of data to analyze, data mining could become quite computationally expensive. Therefore, efficiency becomes a major issue. Currently, we are continuing our efforts to identify ways data should be efficiently analyzed in order to provide accurate and effective results.

In our research, we use the Rule Induction Kit (RIK) and Enterprise Data-Miner (EDM) tools (http://www.data-miner.com) to detect and mine attack signatures. The RIK package discovers highly compact decision rules from data, while the EDM software kits implement the data-mining techniques presented in (Weiss, S. and Indurkhya, N., 1997) and includes programs for (a) data preparation (b) data reduction or sampling, and (c) prediction. Our selection of this tool package was based on criteria related to efficiency (speed, especially when it comes to large amounts of data, as is the case with log files), and portability (multiple platforms), as well as extensibility (where the user can compose new methods with the existing building blocks).

Based on this software platform, we are able to create a sophisticated data mining methodology for efficient intrusion detection.

4.4 Security Safeguards

Safeguards are applied to reduce security risk to an acceptable/desirable level. They may be Proactive to prevent security incidents, or Reactive, to protect information when an incidence is detected. In either case, they must be cost effective, difficult to bypass, and with minimal impact on operations. Examples of safeguards are: avoidance (keeping security incidents from occurring, e.g., by removing vulnerabilities), limiting access (e.g., by reducing the number of entry points where attacks may originate), transference (shifting risk to someone else, e.g., via insurance or outsourcing), and mitigation (minimizing the impact of an incidence, e.g., by reducing its scope or improving detection).

One of the major safeguards is to detect and reduce/remove vulnerabilities. The main reasons for existing vulnerabilities are buggy software design and development, or system administration problems. Existence of bugs in software are due to

- programming for security not being generally taught,
- good software engineering processes not being universal, as well as
- existence of legacy code.

The system administration problems are due to inadequate policies and procedures, or the system administrators being too busy with many machines to administer, too many platforms and applications to support, and too many updates and patches to apply.

For the attack examples given in the previous section, we can offer some rather simple safeguards. For attacks that are based on making multiple requests and ignoring the server acknowledgments, such as ICMP Flood and Smurf Attack, and SYN Flooding, one could employ a timer: if the response does not arrive within a reasonable time, the request could be dropped and the resources freed. For attacks that are based on buffer overflow, one could use operating systems written in "safe" languages that perform range checking (like Java). The HTTP GET attack could be prevented by making sure that programs validate the parameters passed to them, and that file permissions are set properly.

5 CONCLUSIONS

We have first set the stage emphasizing the magnitude of the security problem, raising awareness and focusing on the impact of security. We have detailed two attacks: Denial of Service and the HTTP GET attack, and defined the signature of the latter. The application of data mining techniques for detecting attacks was described. The novelty of our approach is in determining the relevance/importance of different log records, defining intelligent signatures, and using efficient data mining techniques. Preliminary results have been encouraging.

There is significant work still to be done, e.g., improving the effectiveness of attack signatures, developing distributed algorithms for detection/prediction, and improving the efficiency of pattern searching. We are currently working on these issues.

REFERENCES

http://www.w3.org/TR/wsdl

Michael J. A. Berry and Gordon Linoff, Data Mining Techniques, Wiley Computer Publishing, 1997

Computer Emergency Response Team/Coordination Center (CERT/CC) at Carnegie Mellon University's Software Engineering Institute, http://www.cert.org/

www.insecure.org/nmap/nmap-fingerprinting-article.html

NIST ITL Bulletin, "Computer attacks: what they are and how to defend against them," May 1999.

CSI, "2002 CSI/FBI Computer Crime and Security Survey," http://www.gocsi.com/.

The SANS Institute (http://www.sans.org/top20/), May 2003

Douglas Comer, Internetworking with TCP/IP Vol. 1: Principles, Protocols, and Architecture (4th Edition), Prentice Hall, 2000

www.insecure.org/sploits/ping-o-death.html

www.w3c.org?

S. McClure, S. Shah, and S. Shah, Web Hacking: Attacks and Defenses, Addison Wesley, 2003

http://www.interwld.com/pico/subs/pico_Environ_IP_Logging.htm

http://packages.debian.org/unstable/net/ippl.html

W. Lee and S. J. Stolfo, "Data Mining Approaches for Intrusion Detection," Usenix Security Symposium, San Antonio, Texas, July 1998

Jeremy Frank, "Artificial Intelligence and Intrusion Detection: Current and Future Directions," June 1994 (http://citeseer.nj.nec.com/frank94artificial.html)

Zhen Liu, German Florez, and Susan Bridges, "A Comparison of Input Representation in Neural Networks: A Case Study in Intrusion Detection," Proc. International Joint Conference on Neural Networks, May 12-17, 2002, Honolulu, Hawaii. http://www.data-miner.com

S. Weiss and N. Indurkhya, Predictive Data Mining: A Practical Guide, Morgan Kaufmann, 1997.

Magnus Almgren, Herve Deba, and Marc Dacier, "A Lightweight Tool for Detecting Web Server Attacks," http://www.ce.chalmers.se/~almgren/Publications/almgren-ndss00.pdf

S. Forrest, S. A. Hofmeyr, A. Somayaji, and T. A. Longstaff, "A Sense of Self for Unix Processes," Proc. 1996 IEEE Symp. Security and Privacy, Los Alamitos, CA, pp. 120-128, 1996

S. A. Hofmeyr, A. Somayaji, and S. Forrest, "Intrusion Detection using Sequences of System Calls," Journal of Computer Security Vol. 6, pp. 151-180, 1998.

TOWARDS AN ALTERNATIVE WAY OF VERIFYING PROXY OBJECTS IN JINI

Nikolaos Papamichail and Luminita Vasiu
School of Computer Science, Middlesex University, London, UK
Email: n.papamichail@mdx.ac.uk, l.vasiu@mdx.ac.uk

Keywords: Jini Security, Proxy Trust Verification.

Abstract: Jini networking technology represents an exciting paradigm in distributed systems. Its elegant approach in computer networking possesses immense advantages, but also generates security problems. Extensive research has been undertaken and existing security methodologies have been applied to provide a safe execution environment. However the unique nature of Jini has made it hard for traditional security mechanisms to be applied effectively. Part of the problem lies within the downloaded code and in the lack of centralised control. Current solutions are based on assumptions; therefore they are inadequate for enforcing the security requirements of the system. The goal of our research is to increase the security of the Jini model without altering its initial characteristics. We present our preliminary research efforts in providing an alternative, fault tolerant security architecture that uses a trusted local verifier in order to evaluate and certify the correctness of remote calls.

1 INTRODUCTION

Jini networking technology (Sun Microsystems Inc.2003a; http://www.jini.org/) presents an exciting paradigm in distributed computing. Based on the Java programming language, it allows the development of spontaneous networked systems. Users and applications are able to dynamically locate one another and form on-the-fly communities. Unlike traditional systems that rely on a fixed protocol and central administration, Jini requires no further human intervention once being set up. It employs strong fault-tolerance mechanisms that do not attempt to eliminate or hide the fact that network failures may happen. On the contrary it provides a programming model and an infrastructure that allow developers to recognise and isolate any faults that might occur.

When Jini was made publicly available, no security has been taken into consideration. The Java language alone was not adequate to cope with the security required in a distributed setting. Although some solutions have been proposed, Jini lacked a generic security model that could be applied to counter any threats that might arise. The Davis project (http://davis.jini.org/) presents such a security model that has been recently incorporated into the latest Jini release. The security model is based on well known and proven techniques to enforce the basic requirements for network security. However, some of the mechanisms that Jini employs are unique in distributed computing. Additionally, neither any real world applications that make use of the model nor a formal evaluation of it have appeared yet. Thus any assumptions about the correctness of the design and the degree of security provided might prove to be mistaken. The purpose of our research is to examine the security model employed by Jini technology for any potential security faults and propose appropriate modifications. In this paper we focus in the algorithm responsible for verifying trust in Jini proxy objects.

The rest of the paper is organised as follows. Section 2 presents an overview of the Jini programming model and infrastructure, particularly the components that constitute a Jini system and other mechanisms relevant to Jini operation. Section 3 presents some security problems related to proxy objects, Lookup Services and Jini Services while Section 4 presents an overview of the current Jini security model, the Davis Project, and a critical approach to its proxy verification algorithm. Section 5 presents an outline of two proposed solutions to the issues related with proxy object verification and the advantages that they may possess. Section 6 presents related work and some concluding thoughts are drawn in Section 7.

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 11–18.

2 BACKGROUND

Jini (Sun Microsystems Inc., 2003a; http://www.jini.org/) is a distributed system based in Java that allows the establishment of spontaneous network communities or federations. To make that possible, Jini provides the following:

An infrastructure that enables devices, human users and applications to dynamically discover one another without any prior knowledge of their location or of the network's topology and form dynamic distributed systems. The infrastructure is composed of a set of components based on Jini's programming model. Parts of the infrastructure are the discovery join and lookup protocols and the Lookup Service (Sun Microsystems Inc., 2003a). A programming model that is used by the infrastructure as well as by services. Besides service construction, the programming model provides interfaces for performing leasing as well as event and transaction handling.

Services that are employed inside a federation and provide some functionality. Services exploit the underlying infrastructure and are implemented using the programming model.

2.1 Services

Every entity that participates in a Jini system and provides some functionality is perceived as a service. No separation is made regarding the type or the characteristics of the service. A service could be either a hardware device, a piece of software or a human user. Jini provides the means for services to form interconnected systems, and each one separately to offer its resources to interested parties or clients. The separation between a service and a client, however, is sometimes blurred, as sometimes a Jini service may act both as a service and a client.

A word process application, for example, is perceived as a service by any human user that writes a document, although the same application acts as a client whenever it uses a device such as a printer. The latter is again a Jini service, thus for the infrastructure the word application is now its client.

2.2 Proxy Objects

In order for services to participate in a Jini system they must create an object that provides the code by which they can be exploited by potential clients, the proxy object. The proxy object contains the knowledge of the service's location and the protocol that the service implements. It also exposes an interface that defines the functions that can be invoked. A client is able to make use of a service only after the correspondent service's proxy object is downloaded to the client's local space. By invoking functions defined in the proxy interface, clients are able to contact and control services. Clients need only to be aware of the interface that the proxy implements and not of any details of the proxy implementation.

2.3 Lookup Service

The Lookup Service (LUS) is a special kind of service that is part of the Jini infrastructure. It provides a mechanism for services to participate in a Jini system and for clients to find and employ these services. The Lookup Service may be perceived as a directory that lists all the available services at any given time inside a Jini community. Rather than listing String based entries that point back to the location of a service, the Lookup Service stores proxy objects registered by Jini services.

2.4 Discovery Join and Lookup

Relevant to the use of Lookup services are three protocols called discovery, join and lookup (Sun Microsystems Inc. 2003a). Discovery is the process where an entity, whether it would be a service or a client, is trying to obtain references to a lookup service. After a reference has been successfully obtained, the entity might register a proxy object with the Lookup service (join), or search the Lookup Service for a specific type of service (lookup). The discovery protocol provides the way for clients and services to find available Lookup Services in the network, and for Lookup Services to announce their presence.

3 JINI SECURITY ISSUES

Typically security is concerned with ensuring the properties of confidentiality, integrity, authentication and non-repudiation (Menezes et al., 1996):

- Confidentiality ensures that information remains unseen by unauthorised entities
- Integrity addresses the unauthorised alteration of data
- Authentication is the verification of identity of entities and data
- Non-repudiation prevents an entity from denying previous commitments or actions

These properties are generic and apply to a wide variety of systems. Inside Jini, no prior knowledge

of the network's infrastructure is assumed. For that reason, Jini is not only bound to security problems related to distributed systems, but also to any additional issues that the spontaneity of the environment invokes. The following components present different security requirements and they will be examined separately.

3.1 Proxy Object Issues

Nothing should be able to alter the state of the proxy object, either by intention or by fault. That means that the integrity of the proxy object must be ensured (Hasselmeyer et al., 2000a). Since the proxy object is downloaded from an unknown location in the network, neither the source nor the intentions of the proxy object can be verified. Therefore, even the act of downloading the proxy of a service is considered by itself a security risk. Moreover, the proxy is responsible for performing the communication between the client, and the service that the proxy represents. Therefore the integrity and confidentiality of the communication has to be preserved, since the communication link might be intercepted, altered, or simulated by someone with malicious intentions. The privacy and anonymity of the client may be abused, because the client can not be ensured that the proxy does indeed provide the functionality it claims (Hasselmeyer et al., 2000a). On the other hand it has to be verified that any data that needs to be supplied to the proxy object, for the interaction with the service to take place, reaches the appropriate service (JAAS).

3.2 Lookup Service Issues

The Lookup Service lacks any mechanism for authenticating services (Schoch et al., 2001). That means every service can discover the Lookup Service and register its proxy. Malicious proxies may register and pretend they provide some functionality, while they don't. Moreover, every client can search the Lookup Service and find which services are provided. Some services may require only registered users to access them. Therefore access control mechanisms need to be imposed. Additionally, clients might encounter unfairness while searching the Lookup Service for available services (Hasselmeyer et al., 2000a). There is no way a client of a service can be assured that he received the best available service from the Lookup Service. The fact that every service can register and even re-register with the Lookup Service can lead to "man-in-the-middle" attacks (Schoch et al., 2001). A malicious service just has to re-register its proxy with the same service ID as the original one. Every time a client tries to access the required service, the Lookup Service may provide him with the new, malicious proxy. The client is unaware of the change, as the new proxy looks like it implements the same interface as the original one.

3.3 Service Interaction issues

In order for an interaction between two services to take place, the service acting as a client must first locate the provider of the desired service, via the process of discovery, and then download its corresponding proxy object. However, in a spontaneous environment like Jini, hundreds of services may be present at the same time and many of them may provide the same functionality. No standard names or address for recognising individual services exist, besides a unique service ID that is assigned by the Lookup Service. However, it is dependent upon the provider of each service to decide whether or not the assigned ID will be stored and used in any future transactions. Therefore clients have to be able to authenticate the services they access (Eronen et al., 2000). Similarly, the service provider has to be able to authenticate clients that try to use its resources and call its provided functions.

Another aspect in the service interaction is different access levels (Kagal et al., 2001). An obvious solution to this problem is the integration of access control lists. Every user could be identified by a unique username and password that would grant him or deny certain permissions. However, new problems arise, like the distribution of the appropriate keys and the way that the permissions are to be decided.

4 THE DAVIS PROJECT

The Davis project (http://davis.jini.org/) is an effort led by Sun Microsystems' project team responsible for the development of Jini. The purpose is to satisfy the basic Jini requirements for security, by providing a security programming model that would be tightly integrated with the original Jini programming model and infrastructure. Part of the requirements (Scheifler, 2002) has been to avoid changing any existing application code by defining security measures at deployment time. Also to extend the security mechanisms provided by the Java programming language, such as the Java Authentication and Authorisation Service (JAAS).

The Davis project has been integrated with the original release of Jini networking technology (Jini specifications archive – v 2.0) resulting in the

release of the Jini starter kit version 2 (Sun Microsystems Inc., 2003a).

4.1 Constraints

In order to support a broad variety of applications and requirements, the security model dictates that both service providers and their clients should specify the type of security they require before any interaction between them takes place. Decisions upon the type of the desirable security are expressed by a set of constraints that have the form of Java objects. Any service that wishes to incorporate security in its current implementation has to implement a proxy object that implements a well-known interface (Sun Microsystems Inc., 2003b). The interface defines a method for clients and services to set constraints to the proxy object. If the proxy implements that interface, all the imposed constraints apply to every single call through any method defined by the proxy. The basic constraints are the equivalent of Boolean constants that allow decisions upon the type of security required to be specified in proxy objects. Typically service providers specify the constraints during the proxy creation, while clients set the constraints after the proxy object has been downloaded. Constraints specify only what type of security is expected but not how this is implemented.

The security model dictates that constraints imposed by services and clients are combined to a single set of constraints. If any of them contradict with each other then no calls are performed. It is possible, however, that alternative constraints are defined. This provides an elegant way for all parties participating in a Jini interaction to have direct control over the security imposed.

4.2 Object Integrity

There are two mechanisms that the current security model employs to provide integrity for the code of proxy objects. Both assume that the http protocol is used. The first mechanism is http over SSL (https) (Rescorla, 2000), the standard protocol for providing web site security in terms of server authentication, confidentiality and integrity. The other is a custom defined protocol called HTTPMD (Scheifler, 2002; Sun Microsystems Inc., 2003b) The proxy object consists of code which is downloaded by clients, and data which is downloaded from the service. Therefore to ensure total integrity these mechanisms have to apply to both the location where the proxy object is downloaded from and the location of the object's codebase. Along with integrity, the https protocol provides confidentiality and encryption, resulting in additional overhead when these are not required. In these cases the HTTPMD protocol is used. The location of objects, including their code, is specified by a normal http URL. Attached to it is a cryptographic checksum of the contents of the code, a message digest (Rivest, 1992). By computing the checksum of the downloaded data and code and comparing it with the attached message digest, clients are ensured that integrity has been preserved, since any modification in the contents would result in a different message digest.

4.3 Proxy Trust Algorithm

In terms of deciding whether a client trusts a proxy object downloaded by an unknown source, the current model (http://davis.jini.org/) employs the procedure described below. It is assumed that the client has already downloaded a proxy object from somewhere but it can not yet trust neither the proxy object not its correspondent service. Initially the client performs an object graph analysis of the proxy object. By checking recursively all the classes that the object is composed of it can be determined whether these classes are local or not. If the classes are local, in perspective to the client, then the proxy object is considered trusted. This is accepted on the basis that all local code is considered trustworthy. In the case where the proxy object is not fully constructed of local classes, the following components take part in the proxy trust verification algorithm:

1. Proxy object
This is the object that implements the server's functionality. It is downloaded by the client, traditionally from the Lookup Service and it contains the knowledge of how to communicate back with the server. All remote calls to the server are passing through this object and this is the object that needs to be verified.

2. A 'bootstrap' proxy
If the object graph analysis proves that the classes used for the construction of the proxy object are not local relatively to the client, the client uses the initial proxy object to request another object called the 'bootstrap' proxy. The bootstrap proxy should be only consisted of local classes (relevant to the client). The purpose is that clients can trust an object that only uses local code to run. The 'bootstrap' proxy is also used to authenticate the server to the client, as well as to provide him with the verifier described next.

3. A proxy Verifier
The Verifier is an object sent to the client by the

server, using the 'bootstrap' proxy. It checks the downloaded proxy object in order to verify whether the server trusts the initial proxy object or not.

A client obtains a proxy object from the network using Jini discovery and lookup mechanisms. The client examines whether the proxy object is using local code (relative to the client). Since this is normally not the case the client has to verify whether the proxy object can be trusted. The way that this is performed by the current security model is illustrated in Figure 1.

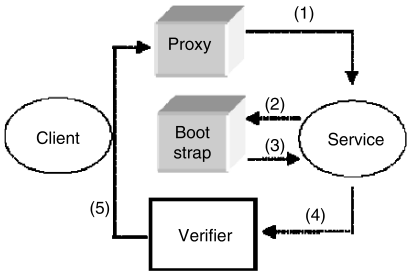


Figure 1: The proxy trust authentication employed by the current security model.

In order to verify that the proxy originates from a legitimate service, the client has to contact the same service and ask the service whether the proxy should be trusted. Since there is no way to directly contact the service, the client places a call through the proxy it can not yet trust, asking for a 'bootstrap' proxy (1). The bootstrap proxy has to use only local classes, relative to the client, in order to be considered trustworthy. After the bootstrap proxy is downloaded to the client's local address space (2), and the locality of the classes that compose the bootstrap object is verified, the client performs a call through it (3). Part of the call is to request from the service to authenticate. After the service has authenticated successfully, it passes a verifier object to the client (4). Finally the verifier is used to verify the legitimacy of the initial proxy object (5).

4.4 Critical Review of the Proxy Trust Algorithm

A number of potential problems might arise from the verification algorithm described above. The first is that clients have to rely on an object downloaded from an unknown source (the proxy object) to obtain the bootstrap proxy. In order for the latter to occur, clients have to place a remote call through an untrusted object. Since the functionality that the proxy object implements is unknown, clients may unintentionally execute an operation that presents a security risk in case the proxy is a malicious object. The second problem is that the service provider has to have some knowledge of the type of classes that are local to the user. If the bootstrap proxy is not consisted entirely by local classes, relevant to the client, the client would not utilize it to obtain the verifier.

A third type of problem is related to the way and type of checks that the verifier performs to the proxy object. There is no standardised set of tests that could be performed, since these are left for the service providers to implement. The method suggested is that the verifier carries the code of the proxy object. By checking the equality of the code that the verifier carries with the code of the proxy object, it is possible for a service to identify the correctness of the proxy object. However, there is no way to ensure whether the checks performed are adequate or if any checks are performed at all.

Therefore a 'lazy' verifier that just confirms the correctness of proxies without performing any checks might incorrectly identify a malicious object as a legitimate one.

Finally faults might occur if a service provider updates the implementation of the proxy object without updating the implementation of the verifier too. In that case legitimate proxy objects would not be able to be identified correctly, since the equality check would fail. Therefore the service provider might unintentionally cause a denial of service attack not initiated by a malicious client, but by himself.

5 AN ALTERNATIVE WAY OF VERIFYING PROXY TRUST

Instead of relying on the untrusted proxy object downloaded from an unknown source to obtain a proxy verifier, clients might be able to protect themselves from malicious proxy objects by using their own local verifier. The verifier is generated locally by clients before any participation in a Jini federation takes place. In order to create the verifier, clients specify their security requirements such as authentication, confidentiality and integrity. These requirements are injected to the verifier and might vary for different scenarios. Specification of the security requirements is similar to the concept of constraints specified by the current Jini security model (http://davis.jini.org/). This permits the specification of application independent security requirements and allows better interoperability with the current security model. The difference is that the

client requirements are not injected into a downloaded proxy, but into the locally generated verifier.

The notion of a locally generated verifier is central to all of the proposed solutions. The operations that the verifier performs, however, are different in every variation of the algorithm. The entities employed in all the solutions proposed here case are the following:

- Client: The entity that wishes to use a service. Clients need to be protected from any potential hazards.
- Proxy object: Typically the object that is downloaded by clients and used to access services. Presents the major source of incoming threats.
- Local Verifier: An entity generated by clients before any interaction with downloaded objects takes place. Used to either verify proxy objects or isolate clients from them.
- Service: The entity that lies somewhere in the network and provides some functionality. Services supply proxy objects and should be considered untrusted.

In every proposed solution it is assumed that a service has already discovered an available Lookup Service and registered its proxy object. The client is ready to perform discovery and lookup in order to obtain a proxy object from the same Lookup Service.

5.1 Proxy Verification Based on a Local Generated Verifier

In order to verify that the downloaded proxy object can be trusted, the following process is performed:

1. Before any discovery process takes place, the client generates a local verifier
2. Client's security requirements are injected to the verifier by the client
3. The client performs discovery of the Lookup Service and downloads a service proxy object
4. Before any interaction with the proxy takes place, the proxy object is passed to the verifier
5. The verifier performs a series of security checks according to the client requirements and makes a decision on behalf of the client about the trustworthiness of the proxy object
6. In case the verifier has decided that the security requirements are satisfied, the client interacts with the service through the proxy object as defined by Jini programming model.

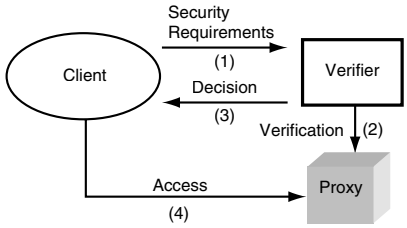


Figure 2: Proxy trust verification by a local verifier.

The described process is illustrated in Figure 2 Initially the client generates the verifier and specifies the security requirements (1). The verifier performs a series of tests to verify trust in the proxy object (2). The result of the verification procedure is expressed as a decision and the client gets notified (3). If proxy has been considered to be trustworthy, the client is allowed to contact the proxy object (4) and access the related service. Comparing this solution with the default proxy verification algorithm, in both algorithms the client is responsible for specifying the type of security required. However, the entity that is responsible for enforcing these requirements is not an untrusted proxy object anymore, but a locally generated verifier. The type of checks performed and the way these are carried out is much more transparent from the client's point of view. Moreover, clients do not have to rely on a verifier object downloaded from a service since the process of such object verifying the initial proxy object is not clear to the client.

Therefore the problem of a service generated verifier that performs no actual check to the proxy object, resulting in the verification of a faulty proxy, is eliminated.

Service providers also do not need to worry about having to provide a bootstrap proxy and a verifier. The only entity that services need to expose is the default proxy object. Absence of a bootstrap proxy eliminates the need for services to implement an object based on the assumption that it would consist of classes that the client already has. Moreover, the current algorithm dictates that every time the implementation of a proxy object changes the verifier object has to change as well, since proxy verification is based on equality checking. Finally by eliminating the need for services to produce two additional objects (the bootstrap proxy and the verifier), administration burden is removed from the service provider.

5.2 Restricting Proxy Object in a Controlled Environment

1. Before any discovery process takes place, the client generates a local verifier
2. Client injects to the verifier the security requirements and the maximum amount of local resources permitted for use by proxy objects
3. The client performs discovery of the Lookup Service and downloads a service proxy object
4. The verifier provides a controlled environment for the proxy object to run. Besides performing security checks to the proxy object, the verifier ensures that the proxy does not use more resources than specified. All requests to and from the proxy object pass through the verifier.

Figure 3 illustrates the followed process. The client generates a local verifier and assigns the security requirements as well as any resources that proxy objects are permitted to use (1). After the proxy object has been downloaded, it is passed to the verifier. The verifier performs similar type of security checks as in proposed solution 1, and additionally provides a controlled environment where proxy objects run. Any client requests and any responses from the proxy object pass through the verifier (2). The same is true for any communication held between the proxy object and its corresponding service.

The advantages of this solution are similar to those of the solution proposed in Section 5.1. The need for service providers to produce additional objects besides the default proxy object is eliminated and so are the assumptions relevant to the locality of classes in the bootstrap proxy and the checks

performed by the service's verifier. Moreover, by restricting execution of the proxy object into a set of finite resources, a protected environment safeguarded by the verifier is created. Verification does not occur only once, but the verifier is monitoring the proxy object continuously. Therefore any potential hazards that might take place during the execution of the proxy are more likely to be identified and get dealt with.

6 RELATED WORK

In (Eronen et al., 2000) certificates are used to establish trust between services and users. Secure interaction is assumed, by allowing users and services to interact only if they carry the appropriate credentials, supplied by a security library. However, these credentials must be assigned to every service of the Jini community before any interaction could be realised. That reduces the spontaneity that Jini provides, and requires prior knowledge of the services' properties to exist, in order for the appropriate permissions to be assigned correctly. Trust establishment is also the purpose of (Hasselmeyer et al., 2000a). Trust establishment is attempted between the Lookup Service, the service provider and the client. The authors propose an extension to the Jini architecture with a certification authority, which provides certifications for the authentication of components. Capability managers are responsible for administering the rights for each user. In that way, different access levels for each client can be easily implemented. Their solution, however, assumes that one central certification authority exists, in order for the appropriate certificates and capabilities to be distributed to every Jini component that exists in the system. Thus, a prior knowledge of every service's characteristics should exist something that is not usually the case in Jini. Moreover, the existence of a centralised authority is opposed to the decentralised nature of the Jini technology. The integration of authorisation and authentication techniques in Jini is also examined in (Schoch et al., 2001). The authors try to achieve that without introducing any additional components, besides the facilities that Jini and Java already provide. They try to prevent man-in-the-middle attacks, by signing the proxy object with a digital signature. This allows the clients to authenticate the source of the provided service, although it still can not be verified how the service users the provided by the service data.

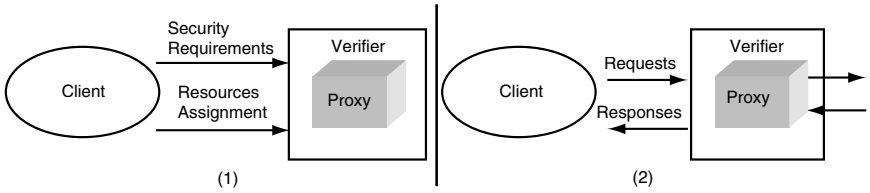


Figure 3: Verifier creation and interaction with the proxy object.

7 CONCLUSION

We presented some security problems related with Jini and how they are countered by the current Jini security model. Our focus is placed in the proxy trust verification algorithm since we believe that an alternative way of verifying proxy object trust might encounter some of the existing problems. We presented our initial ideas in providing an alternative way of ensuring that hostile proxy objects would not impose any risk to clients of the system. We sketched two different approaches in solving the problem. Both involve the concept of a local generated verifier that either verify a downloaded proxy object or impose restrictions to that object's functionality. We also pointed out the advantages of these solutions. Future work includes further rectifying the presented concepts and come up with a viable solution that would integrate with the existing model. Also implement a working prototype and test it in a real world environment.

REFERENCES

Eronen, P., Lehtinen, J., Zitting, J., and Nikander, P., 2000. Extending Jini with Decentralized Trust Management. In Short Paper Proceedings of the 3rd IEEE Conference on Open Architectures and Network Programming (OPENARCH 2000), pages 25-29. Tel Aviv, Israel.

Hasselmeyer, P., Kehr, R., and Voß M. 2000a.Trade-offs in a Secure Jini Service Architecture. In 3rd IFIP/GI International Conference on Trends towards a Universal Service Market (USM 2000), Munich, Germany. Springer Verlag, ISBN 3-540-41024-4, pp. 190-201.

Java Authentication and Authorisation Service (JAAS) http://java.sun.com/products/jaas/ [Accessed 10 Feb. 2004]

Jini specifications archive – v 2.0 http://java.sun.com /products/jini/1_2index.html [Accessed 10 Feb. 2004]

Kagal, L., Finin T. and Peng, Y. 2001. A Delegation Based Model for Distributed Trust. In Proceedings of the IJCAI-01 Workshop on Autonomy, Delegation, and Control: Interacting with Autonomous Agents, pp 73-80, Seattle.

Menezes, A., van Oorschot, P., and Vanstone S. 1996. Handbook of Applied Cryptography. CRC Press. ISBN: 0849385237

Rescorla, E. 2000. HTTP Over TLS, the IETF Network Working Group http://www.ietf.org/rfc/rfc2818.txt [Accessed 09 Feb. 2004]

Rivest, R. 1992. RFC 1321 - The MD5 Message-Digest Algorithm, the IETF Network Working Group, http://www.ietf.org/rfc/rfc1321.txt [Accessed 09 Feb. 2004]

Scheifler, Bob 2002. Comprehensive Network Security for Jini Network Technology Java One Conference Presentation, San Francisco, March 2002 http://servlet.java.sun.com/javaone/sf2002/conf/sessions/display-1171.en.jsp [Accessed 15 Dec. 2003]

Schoch, T., Krone, O., and Federrath, H. 2001. Making Jini Secure. In Proc. 4th International Conference on Electronic Commerce Research, pp. 276-286.

Sun Microsystems Inc. 2003a. Jini architecture specification. http://www..sun.com/software/jini/specs /jini2_0.pdf [Accessed 15 Dec. 2003]

Sun Microsystems Inc. 2003b. Jini architecture specification. http://wwws.sun.com/software/jini/specs /jini2_0.pdf [Accessed 15 Dec. 2003]

http://www.jini.org/ [Accessed 11 Feb. 2004] The Davis project http://davis.jini.org/ [Accessed 11 Feb. 2004

AN EXPERIMENTAL PERFORMANCE ANALYSIS STUDY OF LOSS RATE AND JITTER CHARACTERISTICS IN WIRELESS NETWORKS

M. S. Obaidat[1] and Yulian Wang[2]
[1]*Monmouth University, NJ, USA and* [2]*Tampere University of Technology, Tampere, Finland*
Corresponding Author: Prof. M. S. Obaidat, Department of Computer Science,
Monmouth University, W. Long Brach, NJ07764, USA
E-mail: Obaidat@monmouth.edu

Keywords: Wireless networks, Jitter, loss rate, mobile IP, performance evaluation, bit error rate (BER), Quality of Service (QoS), Diffserv, resource reservation, performance evaluation.

Abstract: Among the challenges in wireless networks is the high bit error rate, which is due mainly to atmospheric noise, physical obstructions found in the signal's path, multipath propagation, interference from other systems and terminal mobility. This high bit error rate makes it difficult to offer guaranteed services over the wireless link. In this paper, we present an experimental analysis study on the loss rate of wireless systems using six different scenarios. We also present an experimental study of jitter for UDP traffic over wireless Mobile IP networks. It is found that in order to provide different Quality of Service, QoS, to downstream traffic flows and control network's loss rate and jitter, it is not enough to have only DiffServ flow control mechanisms. A protocol for wireless link resource reservation and cooperation by senders is also needed. We identify the relationship of jitter, loss rate and class allocation's effect using Class Based Queueing (CBQ) and the packet sending rates in the wireless networks. It is found that loss rate and jitter can be controlled with DiffServ flow control mechanism, but it requires that the total traffic rate should be within the limit of the wireless link capacity. Various tests have been conducted under different settings and operating conditions.

1 INTRODUCTION

1.1 Study of Loss Rate

There are fundamental differences between wireless and wired LANs, which pose difficulties in the design of such systems and protocols (Nicopolitidis, P. et al., 2003), (Nicopolitidis, P. et al., 2002). (The wireless medium is characterized by high bit error rates (BERs) that can be ten times than that for wired LANs. Moreover, errors in wireless LANs occur in bursts, whereas in traditional wired systems errors appear randomly. Among the challenges in wireless networks are: (a) wireless medium unreliability, (b) spectrum use, (c) power management, (d) security, (e) routing, and (f) interfacing with wired networks. The phenomena causing reception errors in wireless systems are: (a) free space path loss, (b) Doppler shift due to station mobility, and (c) multipath propagation due to mechanisms such as reflection, diffraction, and scattering. Such mechanisms cause the signals to travel over many different paths (Nicopolitidis, P. et al., 2003), (Green, D. and Obaidat, M. S., 2003). Mobile IP protocol is an extension of the Internet protocol intended to support mobility in the Internet across all kinds of networks, both fixed (wire-line) and wireless types. When employed with wireless access networks, it can be used to create truly mobile networks. In practice, it enables people to access the Internet continuously with their laptop computers and other portable IP-capable devices while moving around an area covered by wireless LANs.

Mobile IP was developed in response to the increasing use of mobile computers in order to enable them to maintain Internet connection during their movement from one access point to the other. The term mobile here implies that the user is connected to one or more application across the Internet and the access point changes dynamically.

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 19–29.

Clearly, this is different from when a traveler uses his ISP's account to access the Internet from different locations during his trip (Nicopolitidis, P. et al., 2003), (Papadimitriou, G. et al., 2002). Mobile IP is the modification to the standard IP so that it can allow the client to send and receive datagrams no matter where it is attached to the network. The only main security problem using this mechanism is the redirection attacks, which occur when malicious clients give false information to the home agent in the mobile IP network. The home agent is informed that the client has a new care of address. Thus, all IP datagrams addressed to the actual client are redirected to the malicious client (Nicopolitidis, P. et al., 2003).

These days, wireless networks are accessed freely by Mobile Nodes. However, an accounting and charging system starts to be developed. The idea of charging for network access quickly leads to the question of what kind of Quality of Service (QoS) the customer is paying for and how it is ensured. QoS support is naturally needed for the transfer of multimedia streams. In general, the assumption is that real-time data streams will be carried by wireless Mobile IP networks, which are sometimes called 4th generation wireless networks.

Loss rate is an important performance metric that is used to evaluate QoS over wireless links. It is possible to achieve QoS for Mobile IP over wireless link if we can find the causes that increase loss rate. High losses in wireless networks make it difficult to offer guaranteed services over the wireless link. For TCP connections, high loss rate will introduce extra delay in the data transmission. For UDP connections, it will increase the unreliability of datagram delivery.

The current Internet architecture with its best effort service model is inadequate for applications that need various QoS assurances. Two different models have been proposed for Quality of Service (QoS) in the Internet: the Integrated services (Intserv) and the Differentiated services (Diffserv) models. Intserv provides QoS guarantees to individual streams from end to end, while Diffserv provides QoS assurances to a group of applications. Both of these models have been designed to work for wired networks. Hence, new solutions are needed for providing scalable QoS on wireless Mobile IP networks.

Jitter is considered an important metric that is used to evaluate QoS for real-time streaming traffic. Real-time streaming traffic and multimedia synchronization (Wang, C., et al., 1994) require source clock recovery for smooth playback at the destination (Pocher, H., et al., 1999). The cost of clock recovery depends greatly on the ability of the network to control jitter (Varma, S., 1996), (Wright, D.J., 1996). Thus, jitter control over wireless networks for real time applications has been directly connected to the quality of service, QoS, provided to end users.

The Class Based Queueing algorithm was introduced by Jacobson and Floyd in 1995 (Floyd, S. and Jacobson, V., 1995). It was designed to share limited link resources efficiently by different classes. CBQ is supposed to be used in a router where the links are heavily utilized. It allows traffic flows sharing a data link to be guaranteed some amount of bandwidth whenever congestion happens. In addition, CBQ can be used to give priority over other traffics to packets that require low delay. In this way, the link can be shared by multiple data flows yet still can meet the QoS requirements.

The basic rule of CBQ is that each class is given an average rate or a weight of the total bandwidth. Each class gets different amount of the shared resources. Classes can also be designed in a hierarchical structure. Classes on the same level of hierarchy share the whole resource of the parent class. Packets are sent whenever there are resources available in the class to which the packets belong. However, if there are any unused resources in the parent class, the child classes may borrow them. Priority can be given to each class also. Higher priority traffics can be handled ahead of other lower priority traffics.

We put our emphasis on the downlink traffic jitter control on the wireless link a bit more than uplink traffic. The tests of the downlink behavior are more significant for two reasons. First, most flows requiring QoS are likely to be downstream flows from the Internet to the mobile, such as broadcast audio and video. Second, technical solutions for dividing the downlink capacity such as the CBQ exist and can be deployed incrementally. On the other hand, there are some open problems in the uplink direction.

A testbed system running Dynamics hierarchical Mobile IPv4 and Redhat Linux 6.1 with Class Based Queueing (CBQ) on PCs with 2 Mbps Lucent WaveLAN cards was set up for experiments in a real network environment. A description and analysis of the preliminary results from various experiments are presented in this paper.

The first main goal of this experimental study is to identify the relationship of loss rate and sent packets in both downstream and upstream directions and how Class Based Queuing (CBQ) can be used to provide different services for downstream traffic. We present our test results and give analysis on loss rate of wireless Mobile IP.

The second goal of this study is to identify the relationship between jitter and packet sending rate with and without DifferServ flow control

mechanism. We investigate here, the possibility to provide network jitter QoS guarantees with Class Based Queueing (CBQ) link resource sharing method. Our focus area is the wireless link and the Mobile IP Foreign Agent (FA), the bottleneck network and router. Classes based on Class Based Queueing (CBQ) method are designed and implemented for the Foreign Agent.

2 SYSTEM SETUP

In this section, we describe the setup of the test, performance measures, design of CBQ classes, and the tools used for measuring the QoS parameters. We installed a Mobile Ipv4 Home Agent in one PC, Foreign Agent in one laptop PC, and three Mobile Nodes (MNs) in three laptop PCs. Another PC was used as the Correspondent Node (CN). All Mobile Nodes are forced to connect to the FA by using a dynamic tool that comes with dynamic hierarchical Mobile Ipv4 (Mediapoli, 2000), (The Dynamics, 2000). The whole setup is depicted in Figure 1.

The operating system used in all computers was Linux Redhat version 6.1. Two Mbps WaveLAN cards were used for the three Mobile Nodes and Foreign Agent in order to provide wireless connections between the Mobile Nodes and Foreign Agent. The WaveLAN, CBQ, and some other Quality of Service, QoS, support software modules were compiled and loaded to the Linux kernel.

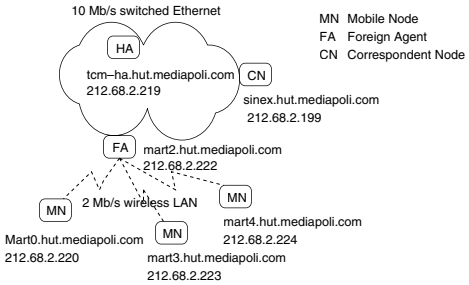


Figure 1: Layout of the test network.

Traffic coming from the CN to MNs went through Foreign Agent. Correspondent Node is connected to Foreign Agent through a high bandwidth Internet connection. Foreign Agent is connected to Mobile Nodes through a wireless connection. In the wired part, the available bandwidth was 10 Mbps. In the wireless part, the available bandwidth was limited to 2 Mbps WaveLAN card. The CBQ was installed in Foreign Agent for downstream traffic. The class design is shown in Figure 2. The bandwidth of the wireless part was divided between two classes. The total bandwidth was set to 1400 kbps. Although the WaveLAN card has a 2 Mbps capacity,

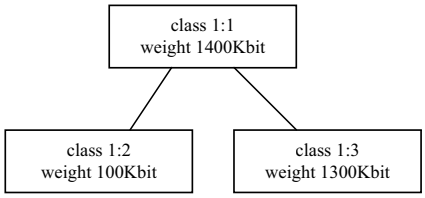


Figure 2: CBQ class design.

the test showed that the total usable bandwidth is 1.4 Mbps. The allocated resource for Class 1 is 100 Kbps with weight 100 Kbps and Class 2 is 1300 Kbps with weight 1300 Kbps. This kind of design of sharing resource is based on the consideration that we can give real time traffic high priority and most of the link resource while give datagram traffic low priority and very little link resource. However, the datagram traffic can use any bandwidth unused by real-time traffic, according to the characteristic of CBQ.

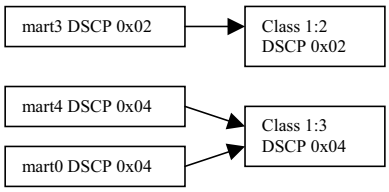


Figure 3: Classifier design.

Packets were marked with different Differentiated Services Code Point (DSCP) values in the IP header before they were sent from the Correspondent Node to Mobile Nodes. In our test system, two DSCP values were used. Thus, all packets were separated into two types; each of which goes through its own class as shown in Figure 3. When these packets come to FA, it will separate them into two classes according to the DSCP values in the IP header of the packets. As shown in Figure 3, DSCP value 0x02 goes to Class 1:2 and 0x04 goes to Class 1:3. The two classes have different bandwidth weights. If the traffic of one class exceeds the bandwidth limit, then the excess packets will be discarded first by CBQ. These two classes can borrow bandwidth from each other if the other has leftover bandwidth. Thus, no bandwidth is wasted.

We installed measurement software called Iperf (Gates, M. and Warshavsky, A., 2000) in the three MNs and CN. Iperf is used to measure the maximum TCP and UDP throughput. It reports bandwidth, jitter (delay variance) and datagram loss. It is a

similar tool to ttcp, but it has overcome some of the limitations of ttcp. Iperf can run for a specified time and can print periodic bandwidth, jitter, and loss reports at specified time intervals. In the following experiments, jitter is calculated in average over 1 second.

3 EXPERIMENTS AND RESULTS

This Section describes the experiments performed and measurement results obtained. We have considered six different scenarios; scenario 1 to 6.

3.1 Scenario 1

In this experiment and the next one, we study the relation between loss rate, jitter characteristics and sending rate in the uplink direction of wireless systems. Three MNs (mart0, mart3, and mart4) are sending data to CN. They are forced to connect to the FA mart2. CBQ is not installed because it doesn't affect uplink traffic in the wireless part. Mart0 sends between 0 seconds to 60 seconds. Mart3 sends between 10 seconds to 60 seconds. Mart4 sends between 30 seconds to 50 seconds. Therefore, we can see only one MN sending, two MNs sending at the same time, and three MNs sending at the same time. Three MNs try to send at a rate of 700 Kbps.

The performance results of loss rate are depicted in Figure 4.

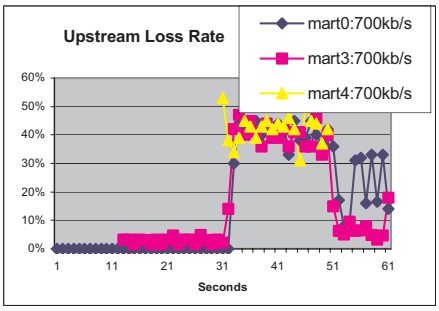


Figure 4: Loss rate chart for scenario 1.

As shown in the figure, we notice that:

1. When only one MN is sending, the total required bandwidth is 700 Kbps which is within the limit of the available bandwidth of the wireless link. The loss rate is about 0.
2. When two MNs are sending at the same time, the total required bandwidth is 1400 Kbps, which is just within the limit of the total available bandwidth. The upstream loss rate for the first stream is almost 0 and for the second stream is about 4%.
3. When three MNs are sending at the same time, the total required bandwidth is 2.1 Mbps, which exceeds the limit of the total available bandwidth. The loss rate for three streams is the same and it is about 42%.

The performance results of Jitter characteristics are depicted in Figure 5.

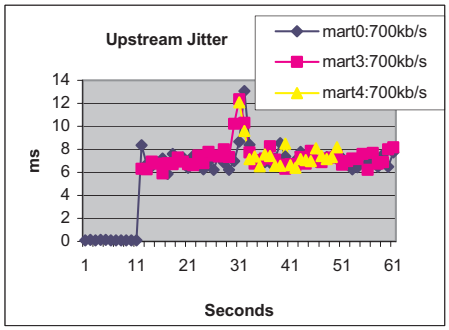


Figure 5: Jitter characteristics for scenario 1.

As shown in the Figure 5, we notice that:

1. Jitter is evenly distributed among different traffic flows.
2. When only one MN is sending, jitter is 0.1ms, which is very small.
3. When two MNs are sending at the same time, jitter is around 7 ms. Jitter is increased from 0.1ms to 7ms when total requested bandwidth increased from 700 Kbps to 1.4 Mbps, which is within the link capacity.
4. When three MNs are sending at the same time, surprisingly the jitter is same as when two MNs are sending. The total requested bandwidth is 2,100 Kbps, which is more than the available bandwidth in the wireless link 1400 Kbps. From our previous related work, we notice that the loss rate is about 42% for all three streams in this situation. This heavy packet loss is mainly caused by data collision and channel contention. We may conclude that whenever we try to observe the jitter characteristic of a stream, we should also consider the packet loss rate as well.

From this experiment, we conclude that jitter characteristic is affected by how many other mobile nodes are sending data at the same time and their sending rate. If the total requested bandwidth is more than the available bandwidth in the link, then we should

consider packet loss rates when we analysis the jitter characteristics. In order to mantain the maintain the traffic flow's jitter at a certain level, some controlling methods must be taken.

3.2 Scenario 2

In this experiment, we considered three MNs that try to send at 1.5 Mbps rate. All other settings are the same as in scenario 1. The purpose of the test is to find out how loss rate and jitter characteristic change when MNs send more UDP packets and require more bandwidth resource than the available bandwidth in the wireless link. Figure 6 summarizes the results obtained in this test.

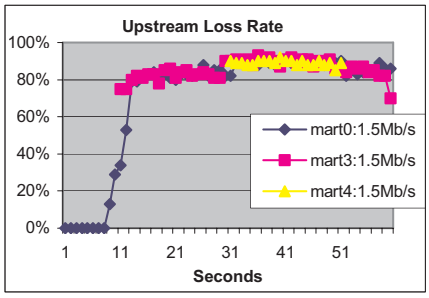


Figure 6: Loss rate chart for scenario 2.

The finding of this experiment can be summarized as follows. When three MNs are sending at the same time, the total required bandwidth is 4.5 Mbps, which is beyond the limit of the total available bandwidth. The loss rate for three streams is same and it is around 90%. We can see that the useful bandwidth is not equal to the available bandwidth 2Mbps in the wireless link. About 90% of bandwidth is wasted due to data collisions. Figure 7 summarizes the results obtained for jitter characteristics in this test.

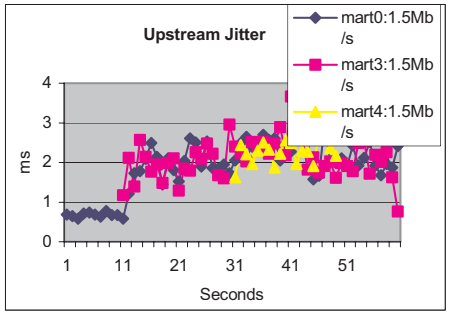


Figure 7: Jitter characteristics for scenario 2.

As shown in the Figure 7, we notice that when two or three MNs are sending at the same time, jitter is around 2.5ms, which is surprisingly small when comparing to the result from scenario 1. From our previous related work, we found that the loss rate is about 90% for all three streams.

We may conclude that whenever we try to observe the jitter characteristic of a stream, we should also consider the packet loss rate. If the packet loss rate is too large, then the jitter characteristic becomes meaningless. There are many studies that have investigated ways to control loss rate and avoid channel contention in the wireless network (Wang, Y. and Obaidat, M. S., 2004), (Kwon, Y. et al., 2003), (Wang, Y. and Obaidat, M. S., 2004).

3.3 Scenario 3

In the following experiments, we study the relation between loss rate, jitter characteristics and sending rate in the downlink direction of wireless Mobile IP system. We also study the impacts of CBQ on the data lose rate in the wireless link.

In this case, we study the downstream bandwidth allocation loss rate and jitter characteristics between three MNs with CBQ method. The CBQ kernel module and the classes shown in Figure 2 and Figure 3 are installed in FA, mart2. Three MNs (mart0, mart3, mart4) are receiving data from CN. The latter sends data to mart3 between 0–90s, to mart4 between 10 and 90s, and to mart0 between 40 and 70s. Mart3 is receiving a stream of 600 Kbps. The packets sent to mart3 are marked with the DSCP value 0x02. Those packets go through Class 1:2. The packets sent to mart0 and mart4 are marked with the DSCP value of 0x04. These packets go through Class 1:3.

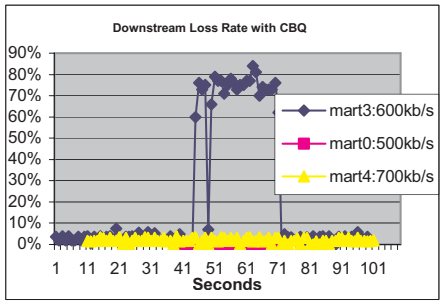


Figure 8: Loss rate chart for scenario 3.

As shown in Figure 8, we can make the following main observations about this experiment:

1. When only one MN is receiving data from the CN, the total required bandwidth is 600 Kbps that is within the limit of the available bandwidth of the wireless link, but beyond the limit of the allocated bandwidth to class 1:2. Since no traffic stream goes in another class, traffic in Class 1:2 can borrow the unused resource in Class 1:3. The loss rate is about 3%.
2. When two MNs are receiving packets and their total requested rate is 1.1Mbps, which is within the available bandwidth, traffic in Class 1:2 can borrow the unused resources in Class 1:3. The loss rate of both streams is around 4%.
3. When three MNs are receiving packets at the same time, the total required bandwidth is 1.8 Mbps, which is beyond the limit of the total available1.4 Mbps bandwidth. The traffic streams toward mart0 and mart4 go through Class 1:3. The total traffic goes through Class 1:3 is 1.2 Mbps that is within the available bandwidth limit. The loss rate for those two streams is the same, which is around 1%.

The stream towards mart3 goes through Class 1:2. The total traffic goes through Class 1:2 is 600 Kbps, which is beyond the limit of the available bandwidth allocated to the class. The loss rate for stream towards mart3 is increased to 79% since it can only borrow 100 Kbps from Class 1:2.

The performance result for jitter characteristics are depicted in Figure 9.

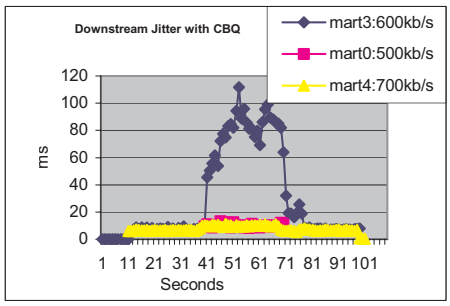


Figure 9: Jitter characteristics for scenario 3.

As shown in Figure 9, we notice that:

1. Jitter characteristics are decided by the allocated resources to each class and the total amount of traffic in the class.
2. When one MN is receiving packets and its requested rate is more than the capacity reserved for its class, but there is enough available bandwidth in the wireless link, the jitter is 0.
3. When two MNs are receiving packets and their total requested bandwidth has been increased to 1300 Kbps that is within the available bandwidth, the jitter increased to 9ms.
4. When three MNs are receiving packets, the jitter for mart4 and mart0 is around 11ms, which has not increased much. Mart4 and mart0 go to the class that has 1300 Kbps capacity, and the total requested bandwidth by them is 1200 Kbps, which is within the class capacity. The jitter for mart3 has been increased from 9ms to 100ms. Mart3 goes to the class that has 100 Kbps capacity and its requested bandwidth is 600 Kbps, which is much more than available bandwidth since now three MNs are receiving at same time. In another class, there is only 100 Kbps unused bandwidth left. Mart3 can borrow this 100 Kbps, but still mart3 cannot get all bandwidth that it needs.

In this experiment, we see that CBQ method can be used to control traffic flow's jitter to a constant level. Mart4 traffic flow's jitter has been controlled at almost constant level. The jitter is not affected by other Mobile Node's sending large amount of data.

3.4 Scenario 4

Contrary to the experiments in scenario 3, experiments in scenario 4 have been performed without CBQ setting. All other settings in this case are the same as in case 3. The purpose of this test is to see the downstream bandwidth allocation between the three MNs without CBQ method. By comparing the results from these two tests, we can analyze the effect of CBQ on the loss rate variations for the traffic towards MNs.

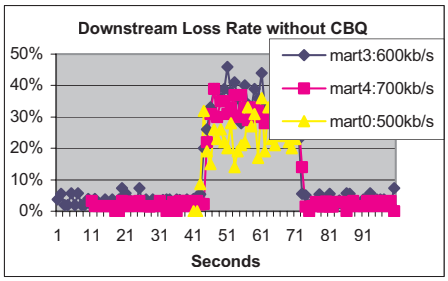


Figure 10: Loss rate chart for scenario 4.

Figure 10 summarizes the results obtained from this experiment. The main findings of this test are: When three MNs are receiving packets at the same time, the total required bandwidth is 1.8Mbps, which is beyond the limit of the total available 1.4 Mbps bandwidth. The loss rate for each of the three traffic

streams is almost same, which is around 30%. Here all traffic flows suffer from the insufficient traffic capacity.

The performance result for jitter characteristics are depicted in Figure 11, which summarizes the results obtained from this experiment.

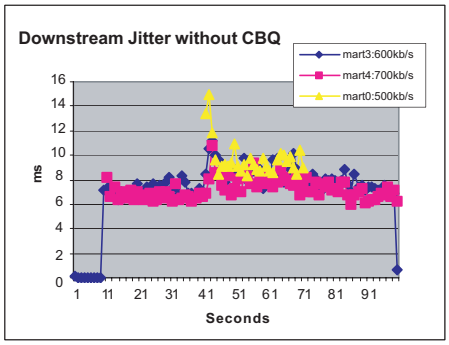


Figure 11: Jitter characteristics for scenario 4.

The main findings of this test are similar to scenario 1 for uplink data. Without CBQ; we cannot see any differentiation of jitter behavior for each traffic flow. All traffic flows jitter are almost same. From our previous work, we found that when three Mobile Nodes are receiving data at same time (the total requested bandwidth is 1800 Kbps), the loss rate was around 40%. In this period, jitter becomes meaningless.

From scenario 3 and 4, we can conclude that CBQ can be used to control jitter to a constant level when the requested bandwidth is within the bandwidth limit of the class. In scenario 5 and 6, we test if CBQ still works if requested bandwidth exceeds the bandwidth limit for the class.

3.5 Scenario 5

In scenarios 5 and 6, all MNs receive large amounts of data that exceed the available bandwidth. Experiments in scenario 5 are performed with CBQ while in scenario 6 without CBQ. The purpose of experiments in this scenario is to see the downstream bandwidth allocation loss rate and jitter characteristic between three MNs under the condition that all MNs receive large volume of data that exceed the available bandwidth in the wireless link. By comparing the results from these two tests, we can analyze the effect of CBQ on the loss rate and jitter characteristic variations for the downlink traffic. The main observations on the results in this scenario are:

1. When one MN, two MNs or three MNs are receiving, the total required bandwidth exceeds the total available bandwidth resource. The loss rate is almost 100% for all of the streams. From

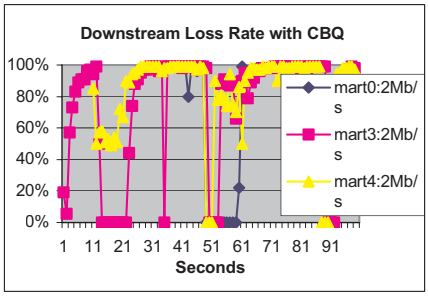


Figure 12: Loss rate chart for scenario 5.

the loss rate chart, we can see that the loss rate is in extremely unreliable. Almost all the bandwidth is wasted due to data collisions.

2. We can conclude that CBQ can not ensure the link quality when the total downstream traffic at FA in each class exceeds the total wireless link capacity allocated to each class.

The performance result for jitter characteristics are depicted in Figure 13.

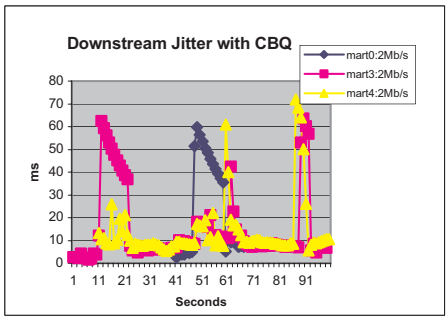


Figure 13: Jitter characteristics for scenario 5.

The main observations on the results in this scenario is that when there are two or three MNs receiving data at the same time, the jitter for each data flow is extremely unstable; it varies between 5 ms to 60 ms. CBQ does not guarantee better jitter characteristics for high priority class. From our previous related work, we noticed that the loss rate is almost 100% for all three streams. The jitter characteristic becomes less important than controlling loss rate in his situation. We conclude that we have to consider loss rate when we analysis jitter characteristics especially in the case when the wireless network resource is in over used condition.

3.6 Scenario 6

Contrary to scenario 5, the test in this scenario has been performed without CBQ setting. All other settings in this test case are same as the test case 5. Figure 14 depicts the downstream loss rate versus time.

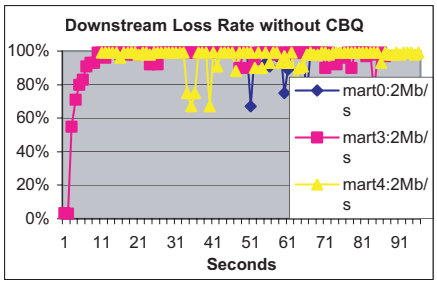


Figure 14: Loss rate chart for scenario 6.

The main observation from the results obtained from this test is that when one MN, two MNs or three MNs are receiving, the total required bandwidth exceeds the total available bandwidth resource. The loss rate is almost 100% for all of the streams.
The performance result for jitter characteristics are depicted in Figure 15.

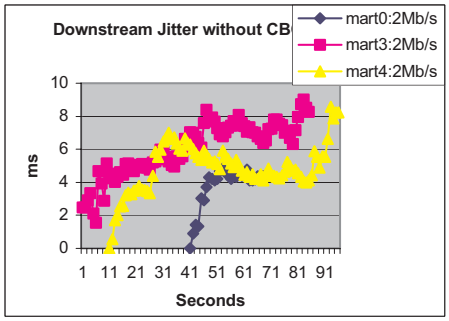


Figure 15: Jitter characteristics for scenario 6.

The result shows that the jitter characteristics become meaningless when all packets almost lost. Again, we conclude here that we have to consider loss rate when we analysis jitter characteristics.

4 DISCUSSIONS

4.1 Loss Rate Analysis

High losses in wireless networks make it more difficult to offer any guaranteed service. The unpredictable losses are mainly due to low quality radio reception and data collisions (Wang, Y. and Obaidat, M. S., 2004), (Claessen, A. et al., 1994). Mobility of Mobile Nodes also increases error rate by approximately 30%. For TCP connections, high loss rate will introduce extra delay in the data transmission. For UDP connections, it will increase the unreliability of datagram delivery. In reference (Xylomenos, G. and Polyzos, G., 1999), the authors point out that some data collisions may sometimes go undetected with WaveLAN so that the error rate visible to higher layers with bidirectional (TCP) traffic increases (Green, D. and Obaidat, M. S., 2003),(Xylomenos, G. and Polyzos, G., 1999), (Nguyen, G. T. et al., 1996).

In wireless LANs with WaveLAN technology, which is used in our testbed, the bandwidth is shared among MNs using CSMA/CA for access control, instead of CDMA/CD. The reason is that it would be expensive to use CDMA/CD because it uses extra bandwidth.

As observed in experiments conducted in scenarios 1 and 2, we found that Wave LAN MAC layer does not divide bandwidth efficiently if the total bandwidth needed by the sending nodes exceeds the link capacity. The loss rate is same for all sending MNs. The bandwidth is divided equally among the sending nodes and there is no technique for allocating different amounts. As mentioned earlier, the unpredictable losses are mainly caused by low quality radio reception and data collisions. The Low radio quality has been improved by enhancing the hardware quality and by technological innovations such as CDMA. Nevertheless, radio transmission is always more prone to errors and link failure than wired networks. Data collisions can be reduced by limiting the total amount of data transmitted by the wireless nodes to the link capacity. There are general two ways to do this: (a) by letting the MN know the total available capacity by broadcasting or by individual signaling, and (b) by letting the MN reserve resources from the nearest FA. The latter keeps accounting of the total available resources. A MN is only allowed to send traffic after it gets confirmation from the FA.

Data collisions can also be avoided by using a token passing protocol that allows only one node to transmit at a time. However, the token-based approach is generally thought to be inefficient. This is due to the fact that in a wireless LAN, token losses are much more likely to happen due to the high bit error rate (losses) of the wireless medium. Moreover, in a token passing network, the token holder needs accurate information about its neighbors and thus of the network topology. Polling, on the other hand, is a more appealing MAC option for a wireless LANs since it offers centralized

supervision of the network nodes. However, constant monitoring of all nodes is required, which is not feasible in the harsh fading environment of a wireless LAN (Wang, Y. and Obaidat, M. S., 2004), (Obaidat, M.S., and Green, D., 2003). Hence, we conclude that a protocol is needed for allocating bandwidth to MNs and MNs must themselves limit their send rate.

In reference (Nicopolitidis, P. et al., 2003), the authors propose a self-adaptive neural-based MAC protocol (SANP) for distributed wireless LANs. According to the proposed protocol, the mobile station that is granted permission to transmit is selected via a neural-based algorithm, which is used to train the system in order to adapt to the network traffic pattern. The neural-like training algorithm utilizes a probability distribution vector, which contains the choice probability for each mobile station in the network. The network feedback plays the role of the system tutor. Following the reception of the feedback after a packet transmission, the neural algorithm performs a simple training procedure in order to reach the goal of "learning" the transmission probabilities of the mobile stations. It was proved that the training algorithm asymptotically assigns to each station a portion of the bandwidth proportional to its needs.

As for the downstream transmission, the current Wireless LAN technology is not mature yet to provide good QoS guarantees. This is due to the high loss (high BER), low bandwidth, Doppler shift due to node's mobility, multiple path propagation, and poor bandwidth characteristic in Mobile IP wireless networks. Moreover, losses in wireless networks occur in bursts, which complicate matters further.

We observed from the experiments in scenarios 3 and 4 that CBQ works well when the total downstream traffic at FA does not exceed the total wireless link capacity. That is, while the amount of traffic goes into some classes do not exceed the allocated resources to the classes, the CBQ works well. From our experiments in scenarios 5 and 6, we observe that CBQ fails when the incoming traffic exceeds the total resources available. That is, CBQ, at least in the tested implementation, cannot deal efficiently with excessive traffic.

Resource reservation is required to ensure quality of data routing over the wireless link. Combined with resource reservation, Class-Based Queueing (CBQ) can provide sufficient QoS to downlink traffic for nodes that do not intentionally exceed their reserved capacity.

To sum up, in order to provide different Quality of Service to downstream traffic flows, it is not enough to have only CBQ implemented in FA. A protocol for wireless link resource reservation and cooperation by the senders are also needed. An example of such a protocol is the one we proposed in reference (Nicopolitidis, P. et al., 2003).

4.2 Jitter Characteristics Analysis

Due to the high loss, mobility and low bandwidth characteristics in Mobile IP wireless networks, QoS is especially difficult to achieve in wireless LANs (Obaidat, M.S., and Green, D., 2003), (Nicopolitidis, P. et al., 2002). From the experiments results, we can see network jitter can be controlled though CBQ queuing method for down link traffics, but it requires that the total data rate for the specific class be within the class capacity. For uplink network jitter control, there is no central control point (such as FA) to install any queuing method to control Mobile Nodes' access to the wireless network. From the experiment results, we can see if the total sending rate of all Mobile Nodes exceeds the wireless link capacity, then the packet loss is very large. To control jitter and packet loss rate, first of all, Mobile Nodes must self-limit the sending rate within the wireless link capacity. Secondly, better MAC layer controlling method for channel contention and collision detection must be used or developed.

From scenario 1 and scenario 2 for uplink stream data experiment results, we can see in general that when there is more traffic in the wireless link, the jitter will become higher. If there is no separate flow control mechanism for each MN, then their jitter values for UDP traffic are same. But surprisingly, when the amount of total traffic sending rate is much higher than the link capacity, the jitter for each data flow is not increasing much. The main reason for this is due to the fact that loss rate is very high in this case.

Mobile Nodes self-limit packet sending rate has significant effect on the QoS guarantee of the jitter, packet loss and link capacity usage. Token bucket can be used to control the packet sending rate and burst size, and leaky bucket can be used to shape the traffic. By using both, Mobile Nodes can effectively control their packets' sending. However, more importantly, MAC layer must have advanced collision detect method to control Mobile Nodes contention for the channel. With current IEEE standard 802.11 for wireless LAN, Media Access Control (MAC) layer uses Carrier-sense multiple access/collision avoidance (CSMA/CA) for collision avoidance. Wireless LANs also use channel reservation techniques by exchanging short "request-to-send" (RTS) and "clear-to-send" (CTS) control packets before the data packet is sent (Wang, Y. and Obaidat, M. S., 2004), (IEEE Standards, 1997). Two major factors affecting the throughput performance in the IEEE 802.11 MAC protocol are transmission

failure and the idle slots due to backoff at each contention period (Kwon, Y. et al., 2003). This protocol is prone to inefficiencies at heavy loads because of higher waste of bandwidth from collisions and backoffs when traffic increases.

To avoid Mobile Nodes from competing for the wireless link resource and control the total traffic rate within the wireless link capacity, we suggest using capacity reservation method. A simple capacity reservation and cancellation protocol is outlined in (Wang, Y. and Obaidat, M. S., 2004), (Wang, Y. and Obaidat, M. S., 2004). Foreign Agent (FA) will be the central control point to maintain the reservation and monitor the traffics. According to the reservation, each Mobile Node limits its traffic rate. When the total traffic sending rate is within the link capacity, we see from our experiments that jitter can be controlled at a constant level.

We observed from scenarios 3 and 4 that network's jitter can be controlled to different levels by the DiffServ flow control mechanism. With CBQ, jitter can be controlled to a constant level as long as the total traffic rate is within the class capacity. If the traffic exceeds the class capacity, then jitter cannot be guaranteed to be at a certain level. It varies according to the total traffic load of the wireless link.

From these experiments, we see that at least one way we can use to control network's jitter; it is to use CBQ method though designing suitable classes for traffic flows and controlling which class each traffic flow goes to. For example, for important real-time traffic, we can let it go to the class with the high link capacity, while for datagram traffic; we can let it go to the class with low link capacity.

It is worth mentioning that in our experiment, we have used static link capacity reservation by giving different fixed weight to each defined class with CBQ. There are some other ways to dividing link capacity, such as dynamic capacity reservation (Braden, R., et al., 1997) and no capacity reservation (Kilkki, K., 1999). Further research is needed in how to use different link capacity reservation schemes to provide QoS guarantees for wireless Mobile IP networks.

From scenarios 5 and 6 experimental results, we observe that high loss rate of traffic (maximum up to 100%) has made jitter characteristic meaningless. Thus, controlling of the loss rate becomes the first priority. CBQ failed to provide QoS guarantee for the downlink traffics. Thus some better mechanism for controlling wireless network access should be provided.

We can conclude that with current 802.11 MAC layer collision control method, resource reservation is required to assure quality of data routing over the wireless link. With resource reservation, we can make sure that the total traffic that goes through the wireless link is within the resources available. Combined with resource reservation, CBQ can provide sufficient QoS to downlink traffic for nodes that do not intentionally exceed their reserved capacity and jitter can be controlled to a guaranteed level. For uplink traffic, self-limit traffic sending rate together with resource reservation can be used to provide sufficient QoS control, and thus jitter can be controlled to a guaranteed level.

5 CONCLUSIONS

To conclude, loss rate is an essential parameter for providing guaranteed service over wireless links. We observed from the experiments presented in scenarios 1 to 6 that the amount of data sent by Mobile Nodes (MNs) is directly related to the packet loss rate. If MNs send more packets than the wireless link resource capacity, then packet loss rate will increase dramatically. To limit the loss rate, we concluded from the upstream sending and down-stream receiving tests that a protocol for allocating resources for MNs is needed. Moreover, MNs must self limit transmission rate according to the reserved capacity. Clearly, a protocol for wireless link resource reservation is needed. Example of such protocols is our recent work presented in reference (Nicopolitidis, P. et al., 2003). This makes the amount of packets sent and received by MNs not exceed the available wireless link resource limit.

In this paper, we also identify the relationship of network jitter and packet sending rate in the wireless Mobile IP networks. The downlink jitter can be controlled with DiffServ flow control mechanism by using CBQ, but it requires that the total traffic rate be within the limit of the wireless link capacity. We propose the use of resource reservation for the wireless link access for both downstream and upstream directions based on the experimental results. In both directions, a Mobile Node should request a reservation from the Foreign Agent, which must keep track of the reservations and enforce them by dropping excess data. The experience and measured results from these experiments were instrumental in identifying the problem areas and the most viable solutions.

ACKNOWLEDGMENTS

The second author would like to thank Helsinki University of Technology Telecommunication and Multimedia lab for the partial support of this work.

REFERENCES

P. Nicopolitidis, M. S. Obaidat, G. I. Papadimitriou and A. S. Pomportsis, Wireless Networks, John Wiley and Sons Ltd., 2003.

M.S. Obaidat and D. Green, "Simulation of Wireless Networks," in Applied Systems Simulation: Methodologies and Applications (M.S. Obaidat and G.I. Papadimitriou, (Eds.), Kluwer, 2003.

P.Nicopolitidis, G.I. Papadimitriou, A.S. Pomports and M.S. Obaidat,' "Self-Adaptive Polling Protocols for Wireless LANs: A Learning-Automata-Based Approach", Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, pp. 3870-3875, Washington DC, October 2003.

G.I. Papadimitriou, M.S. Obaidat, and A.S. Pomportsis, "On the Use of Learning Automata in the Control of Broadcast Network: A Methodology," IEEE Transaction on Systems, Man and Cybernetics-Part B, Vol. 32, No. 6, pp. 781-790, December 2002.

P. Nicopolitidis, M.S. Obaidat, G.I. Papadimitriou and A.S.Pomportsis, "TRAP: a high performance protocol for wireless local area networks", Computer Communications, Elsevier, Vol. 25, July 2002, pp. 1058-1065.

D. Green and M.S. Obaidat, "Dynamic Waveform-Power Adaptation in Mobile 802.11 Wireless LANs," Proceedings of the 2003 International Symposium on Performance Evaluation of Computer and Telecommunication Systems, SPECTS2003, pp. 116-121, Montréal, Canada, July 2003.

Chang-Jia Wang, Liang-Seng Koh, Chao-Hui Wu, and Ming T. Liu, "A Multimedia Synchronization Protocol for ATM Networks", Proc. Intl. Conf. on Distributed Computing Systems, pp. 476-483, Posman, Poland, Jun 1994.

H. Le Pocher, V.C.M. Leung and D. Gillies, "Explicit Delay/Jitter Bounds for Real-time Traffic over Wireless ATM", Computer Networks, Vol. 31, Vo. 9-10, pp. 1029-1048, May 1999.

S. Varma, "MPEG-2 Over ATM: System Design Issues", Proc.COMPCON'96, pp. 26-31, 1996.

D.J Wright, "Voice over ATM: An Evaluation of Implementation Alternatives", IEEE Communication Magazine, Vol. 34, No. 5, pp.72-80, May 1996.

Sally Floyd and Van Jacobson, "Link-sharing and Resource Management Models for Packet Networks", IEEE/ACM Transactions on Networking, 3:365-386, August 1995.

Mediapoli network, 2000. http://www.mediapoli.com.

The Dynamics - HUT Mobile IP System, Helsinki University of Technology, 2000. http://www.cs.hut.fi/Research/Dynamics/.

Mark Gates and Alex Warshavsky, " Iperf version 1.1.1," February 2000. http://dast.nlanr.net/.

Y. Wang, and M. S. Obaidat "An Experimental Analysis Study of Loss Rate in Wireless Mobile IP Systems," Proc. of Applied Telecommunication Sym-posium, ATS2004, pp. 12-17, April 2004 .

Federico Cali, Marco Conti, and Enrico Gregori, "IEEE 802.11 Protocol: Design and Performance Evaluation of an Adaptive Backoff Mechanism," IEEE JSAC, Vol. 18, no. 9, pp. 1774-1786, September 2000.

Younggoo Kwon, Yuguang Fang and Haniph Latchman, "A Novel MAC Protocol with Fast Collision Resolution for Wireless LANs," Proc. of Infocom, 2003. http://www.ieee-infocom.org/2003/papers/21_03.PDF

A. Claessen, L. Monteban, and H. Moelard, "The AT&T GIS WaveLAN Air Interface and Protocol Stack," Proceedings of the 5th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'94), 1994.

George Xylomenos and George C. Polyzos, "Internet Protocol Performance Over Networks with Wireless Links, "IEEE Network, pp. 55-63, July/August 1999.

G.T. Nguyen, R.H. Katz, B.D. Noble, and M. Satyanarayanan, "A Trace-based Approach for Modeling Wireless Channel Behavior, Proc. Winter Simulation Conference, pp. 597-604, Dec. 1996.

IEEE Standards Department, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", IEEE standard 802.11-1997,1997.

22 Robert Braden, Lixia Zhang, Steve Berson, Shai Herzog, and Sugih Jamin, "Resource reservation protocol (RSVP)". RFC 2205, IETF Network Working Group, September 1997.

Kalevi Kilkki, "Differentiated Services for the Internet", Chapter 5.2, pp. 151-155, Macmillan Technical Publishing, 1999.

Y. Wang, and M.S. Obaidat "A Performance Evaluation Study of Jitter Characteristics in Wireless Networks," Proc. of the 2004 Symposium on Performance Evaluation of Computer Systems and Networks, SPECTS2004, pp. 231-237, July 2004.

ON THE SURVIVABILITY OF WDM OPTICAL NETWORKS

Yuanqiu Luo, Pitipatana Sakarindr and Nirwan Ansari
Advanced Networking Laboratory, Department of Electrical and Computer Engineering,
New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA
Email: yl6@njit.edu, ps6@njit.edu, ansari@njit.edu

Keywords: Wavelength division multiplexing (WDM), Network survivability.

Abstract: At a high speed of a few gigabits per second per wavelength, the *wavelength division multiplexing* (WDM) optical networks offer the capacity of several Terabits per second. More bandwidth in each optical channel means more serious loss each time a failure occurs. Therefore, network survivability is a crucial required provision in WDM optical networks. Survivability is the ability of the network to withstand network failures. Many schemes have been proposed to realize the reliable transmission against network failures. This chapter provides an overview of the survivability issue along with the recently developed solutions in WDM optical networks. We classify these schemes based on their optimization objectives, discuss the schemes in each class, compare their strengths and weaknesses, and present the possible future research issues for survivable WDM optical networks.

1 INTRODUCTION

The explosive growth of data traffic poses important emerging bandwidth requirements on today's networks. The large bandwidth of optical fibres in the order of Terabits per second has made the fibres attractive for high-speed networks. The *wavelength division multiplexing* (WDM) technology is playing a major role in the expansion of our networks by dividing the voluminous bandwidth of a fibre into many wavelengths, with each wavelength offering the capacity of a few gigabits per second. As a result of the high volume traffic carried by WDM optical networks, any node or link failure may have severe consequences and could significantly downgrade the services to the worst extent. This is the reason why network survivability is clearly critical to WDM optical networks.

Survivability refers to the network ability to reconfigure or reestablish the traffic transmission upon any failure. The node failure can be a result of the failure of network components such as wavelength cross-connects (WXCs), wavelength transmitters and receivers. The most common link failure is the fibre cut, which may result from the accidental disruption of cables such as construction works, fires, quakes, or even human errors (Ellinas, 2000). Note that a node failure can be decomposed into failures of the links connected to that node, and multiple link failures can be decoupled into several single link failures; most published research as well as this chapter focuses on the single link failure. Many solutions with a variety of optimization criteria have been proposed recently. WDM network survivability issue can be studied from different perspectives, as summarized in Section 2. Our major focus is to classify the representative survivability solutions into three types based on different optimization objectives, and compare their pros and cons as discussed Sections 3, 4, and 5, respectively. The feasible future research directions are proposed in Section 6, and the conclusions are given in Section 7.

2 SURVIVABILITY PERSPECTIVES

Many paradigms have been explored for the optical network survivability. The recovery schemes can be grouped into several types from different viewpoints. This section briefly describes the WDM network recovery schemes from different perspectives. Their performance comparison is also summarized in Table 1.

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 31–40.

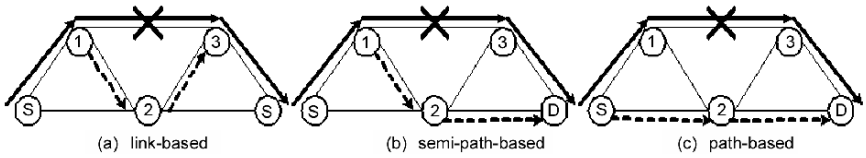


Figure 1: Link-based *vs.* semi-path-based *vs.* path-based recovery schemes.

2.1 Protection *vs.* Restoration

From the viewpoint of resource reservation, existing recovery schemes can be classified as *protection* and *restoration*. In protection schemes, the spare (backup) capacity is reserved during the connection setup and the OXCs and switches are preconfigured. Therefore, the disrupted traffic can be switched to the backup paths (links) as soon as the failure is detected. In restoration schemes, the available spare capacity is allocated to a specific traffic after a failure occurs. While the restoration schemes are more efficient in capacity utilization due to the dynamic sharing of the spare resource, the protection schemes are faster and simpler without additional communications overhead.

2.2 Static Traffic Recovery *vs.* Dynamic Traffic Recovery

Based on the traffic pattern in the network, the recovery schemes can be classified as static and dynamic traffic recovery. In the former, the set of connections is given *a priori*. The recovery schemes configure network equipment and fibre links to minimize the required network resource while providing the recovery against network failures. In the latter, since the traffic changes dynamically, the network configuration should be modified from time to time. To avoid oscillation, a threshold is set to trigger the reconfiguration only when the traffic changes drastically, especially when a network failure occurs. Unlike the static recovery in which the configuration can be done off-line, the dynamic recovery scheme requires fast computation to be done on-line.

2.3 Centralized *vs.* Distributed

From the viewpoint of route computation and fault management mechanisms, there are essentially two paradigms: centralized *vs.* distributed. Centralized recovery schemes depend on the central controller to compute the backup lightpaths and to make the recovery decision based on the up-to-date global network information. Frequent communications between each node and the central controller are required to maintain the accurate link state information. In contrast, distributed recovery schemes make the recovery decision locally. Without the global signalling overhead, the recovery speed is fast. However, each node only has the local information, thus maybe leading to inefficient resource utilization.

2.4 Link-based, Semi-path-based, and Path-based Recovery

Based on the rerouting configuration, the recovery schemes can be grouped as the link, semi-path, and path-based schemes as shown in Figure 1. In the link-based recovery, the single link failure is recovered locally by rerouting traffic around the failed link. Since link recovery is not dependent on specific traffic patterns, it can be preplanned, and therefore fast recovery time can be achieved. In contrast, the failure can be recovered globally by the path recovery. The traffics in the failed link are recovered on an end-to-end basis. All the source-destination nodes of the traffic traversing the failed link reroute the traffic separately and independently. The path recovery scheme requires the involvement of many more nodes and the global network resource information, thus requiring high communications overheads. The semi-path-based scheme is similar to the path-based scheme, except that the disconnected traffic is rerouted from the upstream node of the failed link instead of the source node. Without cranking back to the source node, the semi-path-based scheme has the recovery speed comparable to the link-based scheme.

Table 1: Performance comparison of different recovery perspectives.

Perspective	Type	Recovery Time	Communications Overhead	Resource Utilization Efficiency
Rerouting	Link-based	Fast	Low	High
	Semi-path-based	Medium	Medium	Medium/High
	Path-based	Slow	High	Low
Resource Sharing	Shared	Fast	High	High
	Dedicated	Slow	Low	Low
Fault Management	Centralized	Fast	High	High
	Distributed	Slow	Low	Low
Resource Reservation	Protection	Fast	Low	Low/Medium
	Restoration	Slow/Medium	High	Medium/High
Traffic Pattern	Static-traffic	Fast	Low/Medium	Low/Medium
	Dynamic-traffic	Slow	High	Medium/High

2.5 Shared *vs*. Dedicated Recovery

The recoveries can also be grouped from the viewpoint of resource sharing. In the dedicated recovery scheme, the backup resource is dedicated for a specific primary path (link), and cannot be shared with other backup resource. For shared recovery, several primary paths (links) could share the same backup resource as long as they are disjoint and the failures will not occur simultaneously. Such a sharing results in more efficient resource utilization.

Table 1 summarizes the qualitative performance comparison among different recovery perspectives. In the WDM networks, a particular recovery scheme is essentially a combination of different perspectives. For example, the recovery scheme in reference (Crochat, 1998) is a centralized link protection scheme, and thus it has the properties of fast recovery but with a relatively high communications overhead.

The problem of survivability is basically to optimize the spare network resource in order to realize the reliable functionality against network failures. Therefore, in this chapter, we adopt a new perspective, i.e., the optimization objective, to categorize the survivability schemes into three major classes as design, resource, and traffic optimization recovery schemes. The design optimization recovery is defined as follows: given the network resource and the physical network topology, find the best logical topology that is "immune" from failures, i.e., the logical topology is connected in the event of a single link failure. The resource optimization recovery is defined as follows: given the network topology, find the least network resource required to configure a survivable network against network failures. The resource can be wavelengths, fibres, or wavelength converters. The traffic optimization recovery is defined as follows: given the network topology and specific network resource, find the most guaranteed traffic load or the balanced traffic distribution against failures. Section 3, 4, and 5 will discuss each category one by one. The comparison of strengths and weaknesses among these schemes are summarized in Table 2.

3 DESIGN OPTIMIZATION RECOVERY

We discuss the problems and solutions of the design optimization recovery in this section. Schemes are referred by the authors' names.

The Crochat-Le Boudec scheme (Crochat, 1998) —This scheme optimizes the mapping between the virtual topology and the physical topology to guarantee that each virtual link is independent of others, and no two virtual links share the same physical link. The proposed *disjoint alternate path* (DAP) algorithm maps the virtual topology in such a way that, each virtual link has an alternate virtual path, which shares no physical link with the virtual link itself. In fact, there is a hidden dependency between lightpaths in the virtual topology so that each link is not independent of each other in the physical topology. By deleting the hidden dependencies between primary and backup lightpaths, the DAP algorithm guarantees that when a single link failure occurs, any lightpath can always be recovered by using its predetermined backup lightpath. The complexity is $O(n^4)$, where n is the number of nodes. The studies in (Crochat, 2000) and (Nucci, 2004) extend DAP with the wavelength capacity constraint, and logical topology optimization, respectively. Some of the drawbacks include: first, wavelength converters must be deployed in all nodes; second, the deployed shortest paths may cause an uneven traffic load distribution.

The Modiano-Narula-Tam scheme (Modiano, 2002)—Based on Menger's Theorem, a topology is 2-connected (i.e., redundant) *iff* every cut of the topology has a cut-set size of no less than two, in

which a *cut* is a partition of the set of nodes into two subsets. The edges connecting those two subsets in a cut is called a *cut-set*. The number of edges in a cut-set is the *cut-set size*. The authors proposed the necessary condition for network routing survivability, i.e., none of the physical links are shared by all lightpaths in a cut-set so that any single link failure does not cause a cut disconnected. Based on such condition, a set of integer linear programming (ILP) formulations is presented to solve the following problem: given a physical topology G and the corresponding lightpath requirement matrix X, find the logical topology to route the lightpaths such that the lightpaths are survivable in the event of a single link failure. The applied constraints include lightpath connectivity, lightpath survivability, and physical link capacity. The ILP is further relaxed by enforcing single node to reduce the number of lightpath survivability constraints to n, where n is the number of nodes. Such a relaxation works better as the physical topology becomes denser. Built upon this necessary condition, the authors developed the lower bound on the number of links a physical topology must contain to support lightpath survivability (Narula-Tam, 2004).

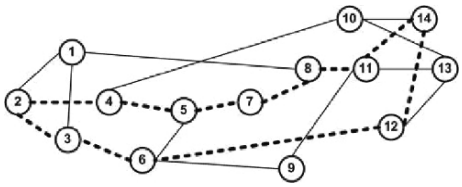


Figure 2: An example of *p-cycle*. Cycle (2, 4, 5, 7, 8, 11, 14, 12, 6, 3, 2) with dotted lines shows an example of the *p-cycle* in the NSFNET.

The Stamatelakis-Grover scheme (Stamatelakis, 2000)—*Stamatelakis* and *Grover* introduced the preconfigured cycles (*p-cycles*) for the survivable network design. *p-cycles* are formed prior to any failure by assigning closed paths in the spare capacity. If a single link in the *p-cycle* fails (called the *on-cycle failure*, e.g., the failure of link (2,4) in Figure 2), the left links in the *p-cycle* form a protection path around that link, exactly working like the rings in SONET. This scheme is particularly effective in mitigating the *straddling* failure (a link which does not belong to the *p-cycle* but its two end nodes that are in the *p-cycle* fails, e.g., the failure of link (5,6) in Figure 2). The *p-cycle* has two restoration paths for a straddling failure. Formed into a closed path, the *p-cycle* provides the recovery speed of rings since each failure only includes the two nodes of the failed link for the recovery operation. By using the spare capacity, the *p-cycle* offers higher capacity utilization since each *p-cycle* contributes to the restoration of more single link failures than a ring. The *p-cycle* approach is in effect a hybrid ring scheme, mixing path restoration for the *straddling* failure with ring recovery for the *on-cycle* failure.

The Medard et al. scheme (Medard, 2002)—The idea of loop-back recovery is realized by assigning two digraphs (directed graphs) for an optical network. The primary digraph is backed up by the secondary digraph. Upon the failure of a link in the primary digraph, the disconnected traffic is carried using the secondary digraph by loop-back. Given an edge-redundant undirected graph $G(N, E)$, this scheme constructs a directed spanning subgraph $B = (N, A)$ and its reversal $R = \underline{B} = (N, \underline{A})$, where each link $a \in A$ in B is reversed to link $\underline{a} \in \underline{A}$ in R. Since B and R are connected, respectively, there exists a directed path in both of them for each pair of nodes. Each of the two conjugated digraphs, B and R, could provide the primary working paths, with the other offering the backup capacity. When a link fails, the disconnected traffic loops back in the secondary digraph to travel around that link.

The Ellinas et al. scheme (Ellinas, 2000)—*Ellinas et al*. proposed the *protection cycle*s for any link failure in a 2-connected digraph. The scheme sets up a double-cycle ring coverage for network G, so that each edge is covered by two cycles. Each cycle works as a primary or a secondary ring. When a link fails, the *automatic protection switching* (APS) mechanism switches the affected traffic into the secondary cycle. The set of the secondary cycles is referred to as the *protection cycle*. For a planar network, *protection cycle* is created by embedding the graph G in the plane and assigning certain directions for the faces. For a non-planar network, *protection cycle* is created by the heuristic *orientable cycle double cover* (OCDC) algorithm. The OCDC algorithm starts at an arbitrary node by adding outgoing edges that satisfy the double-cycle coverage constraint; all of the edges could be covered twice by two different cycles. The APS mechanism is implemented with the cooperation of protection fibres and protection switches. When a fibre link fails, the failure is detected and the protection switches switch the traffic from the primary fibre to the protection fibre. Since only the end nodes of the failed link are involved in traffic switching, *protection cycles* can be configured distributively to improve the recovery speed.

Summary—The *Crochat-Le Boudec* scheme employs DAP to ensure the network connectivity after a failure occurs. It may yield low performance in large networks due to its high complexity. Unlike the *Crochat-Le Boudec* scheme that uses the Tabu

search, the *Modiano-Narula-Tam* scheme formulates the network survivability problem by ILP. It provides the optimal solution for various network topologies. The *p-cycles* in the *Stamatelakis-Grover* scheme are a hybrid approach, which combines path restoration and ring protection. It offers a solution with the ring-like recovery speed and the mesh-like bandwidth efficiency. The *Medard et al*. scheme offers a polynomial time solution based on the generalized loop-back, and is applicable for different network topologies, such as planar, non-planar, and Eulerian networks. On the other hand, the *Ellinas et al*. scheme offers a double cycle cover mechanism, which provides a polynomial time solution for planar networks.

4 RESOURCE OPTIMIZATION RECOVERY

In this section, we discuss the recovery schemes for WDM optical networks that optimize network resource utilization. Such a problem is defined as finding the least network resource assignment against any single link failure for a given network $G(N,E)$. The schemes are referred by the authors' names.

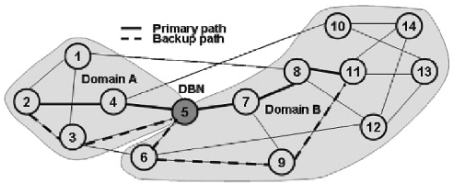


Figure 3: An example of the *Ou et al.* scheme. Lightpath (2, 4, 5, 7, 8, 11) is divided into two primary subpaths (2, 4, 5) and (5, 7, 8, 11) by Domain A and B, respectively. Backup subpath (2, 3, 5) is for primary subpath (2, 4, 5), and backup subpath (5, 6, 9, 11) is for primary subpath (5, 7, 8, 11), respectively.

The Ou et al. scheme (Ou, 2004)—*Ou et al.* proposed to divide an optical network into domains, and a lightpath is thus cut into several subpaths. The shared path protection (SPP) algorithm is then adopted in each domain to provide the least-cost backup subpath for the primary subpath. Two constraints are applied: first, the primary and backup subpaths of an inter-domain lightpath must exit or enter any domain at the same domain-border node (DBN); second, the primary and backup subpaths of an intra-domain lightpath can only use the resource in the same domain. The resource of subpaths is optimized by maximizing backup resource sharing. When a failure occurs, traffic is switched within a domain rather than the entire network, thus contributing to the fast recovery.

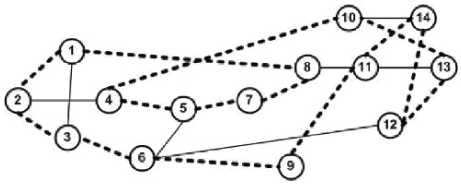


Figure 4: A Hamiltonian cycle with dotted lines in the NSFNET.

The Huang-Copeland scheme (Huang, 2002—A *Hamiltonian cycle* (HC) is a closed path which includes every node in the network exactly once. Figure 4 shows a HC in the NSFNET with dotted lines. Such a cycle is actually the spanning ring of the network. When a single link fails, its two end nodes switch the disconnected traffic to the HC. The network is by design always connected since a single link failure could only reduce the cycle into a spanning tree. Therefore, the HC guarantees the recovery of the network against any single link failure. Since a HC is a spanning ring of the network, it may have a wider recovery coverage than a *p-cycle*. For example, the single link failure of link (10,14) cannot be recovered by using the *p-cycle* in Figure 2. Such a failure can be protected by using the path (10,13,12,14) as shown in Figure 4. In order to cover all links, more than one *p-cycles* are needed while only one HC is enough, and assigning multiple *p-cycles* complicates the network management. The bottleneck for HC is that not every network contains a HC.

The Su-Su scheme (Su, 2001)—The authors proposed the ILP formulations for both off-line and on-line configuration restoration routing. The objective is to maximize the wavelength sharing among the protection paths. The "bucket"-based link metric is applied to measure the path "width", and they indicate the degree of resource sharing among different failures in a link and a path, respectively. The "bucket" is the number of protection wavelengths reserved in link l for the failure of link k. The wavelengths required being reserved in link l for any single link failure equal to the maximum bucket in that link. The "width" of link l is proportional to its reserved wavelengths h. When $h = 0$, link l is called "exhausted" since it does not reserve any wavelength for protection. The "width" of path p corresponding to the failure of link k is the minimum value of the link width in path p.

Maximizing resource sharing is achieved by choosing the widest path among all candidates. This widest path has the largest bucket, and thus the largest number of wavelengths could be shared for protection. The required new resource is minimized by assigning the least hop path as the primary path and the widest path as the protection path.

The Xu et al. scheme (Xu, 2003)—The proposed scheme adopts the shared risk link group (SRLG) information to strengthen the cross-layer protection in WDM networks. It separates an active path (AP) and backup path (BP) into several active segments (ASs) and backup segments (BSs), respectively. The goal is to protect each AS with its corresponding BS rather than to protect the whole AP. Different from the proposal in (Ou, 2004), overlapping links may exist among ASs. Resource sharing among BSs is maximized to achieve high efficiency. The ILP approach is practical for medium-size networks, and the dynamic programming (DP) approach yields the suboptimal results with polynomial time complexity for large networks. Interested readers are referred to (Lei, 2004), (Zang, 2003), and (Qin, 2003) for detailed discussions on cross-layer recovery.

The Ho-Mouftah scheme (Ho, 2004)—The proposed *optimal self-healing loop allocation* (OSHLA) algorithm dynamically selects cycles from a predefined cycle set to cover a given lightpath. Each cycle is assigned a cost based on its sharable capacity. Dijkstra's algorithm is then employed to find the best cycle allocation among different options, aiming to maximize spare capacity sharing. Because the cycle length dominates the computation complexity, proper cycle length limits for typical networks have been further developed from experiments, thus achieving a compromise between efficiency and complexity.

Summary—Generally, this group is developed from the design optimization recovery schemes with the focus on the network resource optimization. The *Ou et al.* scheme maximizes the backup resource sharing within domains. The Hamiltonian cycle scheme saves more resource than the *p-cycles* scheme by aggregating the backup capacity into a spanning ring with the least number of links. The *Su-Su* scheme applies the bucket model and the widest path to ensure that the reserved bandwidth can be maximally shared among multiple link failures as long as they do not occur simultaneously. The *Xu et al*. scheme protects several working segments instead of the whole working path, that results in higher resource sharing than traditional path-based recovery schemes. The *Ho-Mouftah* scheme enhances the SONET self-healing ring by accommodating on-line sharable resource information, and the employment of cycle length limit contributes to its on-line performance.

5 TRAFFIC OPTIMIZATION RECOVERY

The traffic optimization can be divided into two major types: balancing traffic load and maximizing carried traffic load. The two types of traffic optimization problems are closely related. From the point of traffic engineering, the traffic load can be balanced by selecting the link with a light load. By circumventing the heavily loaded links, the traffic blocking probability is reduced, and therefore the overall accepted traffic increases.

The Ruan et al. scheme (Ruan, 2004)—The proposed *routing with load balancing heuristics* (RLBH) algorithm adopts load balancing in restorable path computation. A pair of primary and backup paths is set up for a connection request, and the lighter-loaded links are preferred over the heavier-loaded ones. A *critical index* δ is employed as the threshold to specify the link cost. When the number of free channels over link l is more than δ, the corresponding link cost is set to 1; the link cost is set to ∞ if there are no free channels over link l; otherwise, a constant η ($\eta > 1$) is set as the link cost. When computing the backup path, RLBH prunes the links on the primary path, and prefers the links containing sharable backup channels. The Dijkstra's algorithm is then employed to find the least-cost path pair. RLBH needs the network status information to estimate the number of backup channels in a link, and thereby, *interior gateway protocol* (IGP) has been extended with signalling augment.

The Mohan et al. scheme (Mohan, 2001)—A dependable connection (D-connection), including the primary working and the corresponding back-up lightpath, is established upon a connection request. In order to maximize the number of carried D-connections, network resource multiplexing is employed to share links among lightpaths. Different from other proposals, besides resource sharing among backup lightpaths, this scheme also allows resource sharing among a primary and several backup lightpaths to carry more D-connections. The corresponding assumption is that no single link failure will cause two primary paths to compete for the same backup resource. The computational complexity is O(knw) when the two lightpaths in a D-connection use the same wavelength, and O(k2nw) when they use different wavelengths, where k is the predetermined maximum number of candidate routes for any node pair, n is the number of nodes, and w is the number of wavelengths per fibre.

The Sahasrabuddhe et al. scheme (Sahasrabuddhe, 2002)—The fault management mechanism is proposed to maximize the guaranteed traffic in a network. In the WDM layer, the modified path

protection scheme configures a backup lightpath for each primary lightpath. In the IP layer, the modified restoration scheme also pre-configures the network so that the disconnected traffic can be rerouted over the spare capacity. These two schemes aim to maximize the scalar load factor, α. The whole network traffic variation is modeled as αT, where T is the traffic matrix. The larger the load factor, the more guaranteed traffic can be carried by the network. The heuristic algorithm maximizes α by iterating between two steps: step one attempts to free as many wavelengths and lightpaths as possible while maintaining the load in the maximally loaded link; step two attempts to set up as many guaranteed lightpaths among the freed resource as possible while decreasing the load in the maximally loaded link. To improve the efficiency, the shared-path protection is employed to allow resources shared among different backup lightpaths. The WDM layer shared-path protection offers more guaranteed traffic and much faster recovery time than those of the IP layer restoration, which takes longer time for processing link state updates and recomputing routing tables.

The Qiao-Xu scheme (Qiao, 2002)—The distributed partial information management (DPIM) algorithm protects the traffic from the link failure based on a shared path protection approach. It determines a pair of active path (AP) and backup path (BP) for each bandwidth-guaranteed connection to maximize the number of carried connections. The sum of active bandwidth (ABW) and backup bandwidth (BBW) allocated to all connections is called the total bandwidth (TBW). Each candidate link on a new connection distributedly estimates additional BBW, and the path assignment with the least BBW is selected. Besides connection establishment, DPIM takes connection release into consideration. The overhead of link state information distribution increases the accuracy of BBW estimation, thus maximizing the carried traffic load.

Summary—Load balancing is facilitated in the *Ruan et al*. scheme by directing new traffic to the lighter-loaded links, and the performance is determined by the threshold value δ. The *Mohan et al.* scheme maximizes the carried traffic (i.e., D-connections) by sharing spare wavelength channels among different backup lightpaths or among a primary lightpath and several backup lightpaths. The *Sahasrabuddhe et al*. scheme optimizes the guaranteed traffic and the recovery time by maximizing the *load factor*, α, and taking the advantage of the integration of IP restoration and WDM protection. The *Qiao-Xu* scheme maximizes the bandwidth-guaranteed traffic by configuring the shared protection paths based on the distributed information. Unlike the *Su-Su* scheme [8] that every edge node maintains partial and aggregated $O(E^2)$ network information, the *Qiao-Xu* scheme distributes information around the network, and the information maintained by each node is $O(E)$, where E is he number of links.

6 FUTURE DIRECTIONS

Future research of network survivability research should address the following issues.

6.1 Finer Granularity

Most of the design optimization recovery schemes in Section 3, 4, and 5 consider the recovery in the fibre granularity level, in which the traffic carried by a fibre is backed up by another fibre, and the number of available wavelengths is assumed sufficient for the sake of simplicity. All wavelengths in the backup fibre are reserved in advance or reconfigured in real time. Therefore, the wavelength utilization efficiency is significantly deteriorated. Finer granularity, such as wavelength, should be adopted to improve the efficiency. One possible solution could be extending the current schemes by designing the lightpaths in the wavelength-based layered graph (Luo, 2003) instead of the link-based graph.

Since the backup resource assignment is done in the wavelength level instead of the fibre level, less resource will be reserved, and thus more traffic could be carried.

6.2 Complexity Relaxation

The LP formulations are generally used to provide a mathematical formulation of the network survivability. Owing to the large number of constraints and the network size, solving such a set of formulations is time consuming. Thus, most of the schemes can only be implemented off-line, and appropriate heuristic algorithms are desperately needed to simplify the computation. LP relaxation (Krishnaswamy, 2001) and Lagrangian relaxation (Lee, 2004), (Zhang, 2004) could reduce the complexity. LP relaxation converts the LP problem into the ILP problem by quantizing continuous variables into discrete variables (integers). Branch and bound method is implemented into such a relaxed problem to search for the solution. Lagrangian relaxation approach decomposes the larger multiple constraint LP problem into smaller sub-problems. By relaxing wavelength-related constraints through the use of Lagrange multipliers,

Table 2: Strengths and weaknesses of the recovery schemes.
Note: C = Centralized, D = Distributed, L = Link-based, P = Path-based, R = Resource, DE = Design, T = Traffic

Algorithm	Fault Management	Rerouting strategy	Optimization Strategy	Strengths	Weaknesses
Crochat-Le Boudec (Crochat, 1998)	C	L	DE	• Guaranteed link failure recovery	• Unbalanced traffic load • Full-range wavelength converters are required
Modiano-Narula-Tam (Modiano, 2002)	C	L	DE	• Extendable for multiple failures	• Wavelength continuity and wavelength capacity are not taken into consideration
Stamatelakis –Grover (Stamatelakis, 2000)	C	L	DE	• Similar recovery speed to and higher utilization than rings	• More than one p-cycles are required to cover one network
Medard et al. (Medard, 2002)	C	L	DE	• Polynomial time complexity	• Full-range wavelength converters are required
Ellinas et al. (Ellinas, 2000)	D	L	DE	• High recovery speed	• Protection cycles may not be found in non-planar and Eulerian networks
Ou et al. (Ou, 2004)	D	P	R	• Scalable for large networks	• DBN must be capable of wavelength conversion
Huang-Copeland (Huang, 2002)	C	L	R	• Applicable for diverse traffic granularities	• A Hamiltonian cycle may not exist in an arbitrary network
Su-Su (Su, 2001)	D	P	R	• Balanced traffic load	• Needs to distribute information of wavelength availability
Xu et al. (Xu, 2003)	D	L	R	• Polynomial time complexity	• Wavelength continuity may not be satisfied among path segments
Ho-Mouftah (Ho, 2004)	C	P	R	• Variable loop size	• Needs signalling protocol extensions
Ruan et. al (Ruan, 2004)	D	P	T	• Balanced traffic load	• Full-range wavelength converters are required at all nodes
Mohan et al. (Mohan, 2001)	C	P	T	• Sharing resource among a primary and several backup lightpaths	• Low recovery efficiency
Sahasrabuddhe et al. (Sahasrabuddhe, 2002)	D	P	T	• Interoperability of multiple layers	• The IP restoration scheme may not find an optimal solution because of limited transceivers per node
Qiao-Xu (Qiao, 2002)	D	P	T	• Polynomial time complexity	• Needs accurate link state information

the total number of constraints is greatly reduced. The optimal Lagrange multipliers provide the best tradeoff between the recovery coverage and the resource utilization efficiency. Moreover, the relaxation process must ensure that the resulting solutions are also the solutions to the original unrelaxed problem.

6.3 Fault Recovery in Multifibre Networks

The real practice of installing bundles of multiple fibres motivates the research on the fault tolerance problem for multifibre optical networks. In such a network, a link between two nodes contains several fibres; each supports tens of wavelength channels. If the same wavelength on the next hop is not available, traffic can be switched to another fibre, where the same wavelength is unoccupied. An efficient way to analyze the survivability is based on the layered graph. With each edge representing a specific wavelength channel, and each layer representing the connections in one wavelength, the primary and backup lightpaths can be assigned simultaneously and optimized jointly. An important issue for survivable multifibre WDM networks is to determine whether the increased number of fibres trades off favourably with improved survivability. The research in (Luo, 2004) shows that multifibre WDM networks low the traffic blocking probability. Therefore, it is possible to use fewer wavelengths in each fibre with multiple fibres than with a single fibre to carry the same traffic load.

6.4 Multiple Failures Recovery

Most schemes tackle the single link failure in WDM optical networks. Such a single-link failure model assumes that at most one link can fail at any time, and failures do not occur simultaneously. When a link fails, all links that have failed earlier have been repaired (Mohan, 2001). In fact, with the growth of networks, multiple failures are possible. For example, a construction work may cut a buried optical cable, which has a bundle of fibres, thus leading to several link failures. Moreover, the time to repair a cable may be several hours or days. It is possible that another failure occurs during that interval. To recover multiple failures, the network must be configured with redundancy. In order to protect double failures, the graph must be 3-connected (i.e., it takes the removal of at least three links to disconnect the graph) (Choi, 2002), and several backup lightpaths for a primary lightpath have to be predetermined. Careful configuration of network spare resource must be done to ensure that even when multiple links fail, one in the primary working lightpath and another in the primary backup lightpath, the traffic can be continued through the secondary backup lightpath.

7 CONCLUSIONS

Survivability is a crucial network function for the high-speed WDM optical networks. It seeks to recover network failures by means of the efficient use of spare network resource. Based on different optimization criteria, the existing recovery schemes can be divided into three classes: design, resource, and traffic optimization recovery. Design optimization recovery schemes predesign the whole network and reserve the spare resource for the single fibre failure recovery. Since the network traffic matrix is unknown, such predesign is usually done at the fibre-based level. The design in the wavelength-based level should be employed in order to improve efficiency. Resource optimization recovery schemes minimize the resource used for failure recovery by sharing the spare resource among primary or backup lightpaths. Traffic optimization recovery schemes combine the failure recovery and the traffic engineering. In addition to the recovery provisioning, the carried traffic and load balancing among links are considered. The appropriate relaxation methods could simplify the complexity of the above optimization issues. We have summarized the evaluation of various schemes in Table II, and provided directions for future research on survivability of WDM optical networks.

REFERENCES

Choi, H., Subramaniam, S., and Choi, H.-A., 2002. On double-link failure recovery in WDM optical networks. *Proc. INFOCOM'2002*, vol. 2, pp. 808-816, 2002.

Crochat, O., and Le Boudec, J.-Y., 1998. Design protection for WDM optical networks. *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 7, pp. 1158-1165, Sep. 1998.

Crochat, O., Le Boudec, J.-Y., and Gerstel, O., 2000. Protection interoperability for WDM optical networks. *IEEE/ACM Transactions on Networking*, vol. 8, no. 3, pp. 384-395, June 2000.

Ellinas, G., Hailemariam, A., and Stern, T. E., 2000. Protection cycles in mesh WDM networks. *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 1924-1937, Oct. 2000.

Ho, P.H., and Mouftah, H.T., 2004. A novel survivable routing algorithm for shared segment protection in mesh WDM networks with partial wavelength conversion. *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 8, pp. 1548-1560, Oct. 2004.

Huang, H., and Copeland, J., 2002. A series of Hamiltonian cycle-based solutions to provide simple and scalable mesh optical network resilience. *IEEE Communications Magazine*, vol. 40, no. 11, pp. 46-51, Nov. 2002.

Krishnaswamy, R.M., and Sivarajan, K.N., 2001. Algorithms for routing and wavelength assignment based on solutions of LP-relaxations. *IEEE Communications Letters*, vol. 5, no. 10, pp. 435-437, Oct. 2001.

Lee, S.S. W., Yuang, M.C., Tien, P.L., and Lin, S.H., 2004. A Lagrangean relaxation-based approach for routing and wavelength assignment in multigranularity optical WDM networks. *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 9, pp. 1741-1751, Nov. 2004.

Lei, L., Liu, A., and Ji, Y., 2004. A joint resilience scheme with interlayer backup resource sharing in IP over WDM networks. *IEEE Communications Magazine*, vol. 42, no. 1, pp. 78-84, Jan. 2004.

Luo, Y., and Ansari, N., 2003. Performance evaluation of survivable multifibre WDM networks. *Proc. GLOBECOM'2003*, vol. 5, pp. 2524-2528, Dec. 2003.

Luo, Y., and Ansari, N., 2004. A computational model for estimating blocking probabilities of multifibre WDM optical networks. *IEEE Communications Letters*, vol. 8, no. 1, pp. 60-62, 2004.

Medard, M., Barry, R.A., Finn, S.G., He, W., and Lumetta, S.S., 2002. Generalized loop-back recovery in optical mesh networks. *IEEE/ACM Transactions on Networking*, vol. 10, no. 1, pp. 153-164, Feb 2002.

Modiano, E., and Narula-Tam, A., 2002. Survivable lightpath routing: a new approach to the design of WDM-based networks. *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 4, pp. 800-809, May 2002.

Mohan, G., Siva Ram Murthy, C., and Somani, A.K., 2001. Efficient algorithms for routing dependable connections in WDM optical networks. *IEEE/ACM Transactions on Networking*, vol. 9, no. 5, pp. 553-566, Oct. 2001.

Narula-Tam, A., Modiano, E., and Brzezinski, A., 2004. Physical topology design for survivable routing of logical rings in WDM-based networks. *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 8, pp. 1525-1538, Oct. 2004.

Nucci, A., Sansò, B., Crainic, T.G., Leonardi, E., and Marsan, M. A., 2004. On the design of fault-tolerant logical topologies in wavelength-routed packet networks. *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 9, pp. 1884-1894, Nov. 2004.

Ou, C.S., Zang, H., Singhal, N.K., Zhu, K., Sahasrabuddhe, L.H., MacDonald, R.A., and Mukherjee, B., 2004. Subpath protection for scalability and fast recovery in optical WDM mesh networks. *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 9, pp. 1859-1875, Nov. 2004.

Qiao, C., and Xu, D., 2002. Distributed partial information management (DPIM) schemes for survivable networks.1. *Proc. INFOCOM'2002*, vol. 1, pp. 302-311, 2002.

Qin, Y., Mason, L., and Jia, K., 2003. Study on a joint multiple layer restoration scheme for IP over WDM networks. *IEEE Network*, vol. 17, no. 2, pp. 43-48, Mar.-April 2003.

Ruan, L., Luo, H., and Liu, C., 2004. A dynamic routing algorithm with load balancing heuristics for restorable connections in WDM networks. *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 9, pp. 1823-1829, Nov. 2004.

Sahasrabuddhe, L., Ramamurthy, S., and Mukherjee, B., 2002. Fault management in IP-over-WDM networks: WDM protection versus IP restoration. *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 1, pp. 21-33, Jan. 2002.

Stamatelakis, D., and Grover, W.D., 2000. Theoretical underpinnings for the efficiency of restorable networks using preconfigured cycles ("p-cycels"). *IEEE Transactions on Communications*, vol. 48, no. 8, pp. 1262-1265, Aug. 2000.

Su, X., and Su, C.-F., 2001. An online distributed protection algorithm in WDM networks. *Proc. ICC'2001*, vol. 5, pp. 1571-1575, 2001.

Xu, D., Xiong, Y., Qiao, C., and Li , G., 2003. Trap avoidance and protection schemes in networks with shared risk link groups. *Journal of Lightwave Technology*, vol. 21, no. 11, pp. 2683-2693, Nov. 2003.

Zang, H., Ou, C., and Mukherjee, B., 2003. Path-protection routing and wavelength assignment (RWA) in WDM mesh networks under duct-layer constraints. *IEEE/ACM Transactions on Networking*, vol. 11, no. 2, pp. 248-258, April 2003.

Zhang, Y., Yang, O., and Liu, H., 2004. A Lagrangean relaxation and subgradient framework for the routing and wavelength assignment problem in WDM networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 9, pp. 1752-1765, Nov. 2004.

SIGMA: A TRANSPORT LAYER MOBILITY MANAGEMENT SCHEME FOR TERRESTRIAL AND SPACE NETWORKS*

Shaojian Fu and Mohammed Atiquzzaman
Telecommunications and Networks Research Lab
School of Computer Science, University of Oklahoma,
Norman, OK 73019-6151, USA
Email: {sfu,atiq}@ou.edu

Keywords: Internet Mobility, Mobility Management, Wireless Networks, Handoff management.

Abstract: Mobile IP has been developed to handle mobility of Internet hosts at the network layer. Mobile IP suffers from a number of drawbacks such as requirement of infrastructure change, high handover latency, high packet loss rate, and conflict with network security solutions. In this paper, we describe the architecture of Seamless IP diversity-based Generalized Mobility Architecture (SIGMA) - a new mobility management scheme. SIGMA utilizes IP diversity to achieve seamless handover, and is designed to solve many of the drawbacks of Mobile IP, including requirement for changes in infrastructure. The survivability and security of SIGMA is evaluated and shown that SIGMA has a higher survivability than Mobile IP - thanks to its centralized location management scheme. SIGMA can interoperate with existing network security infrastructures such as Ingress filtering and IPSec fairly easily. We also show the application of SIGMA to manage satellite handovers in space networks.

1 INTRODUCTION

Mobile IP (MIP) (Perkins, 2002; Perkins, 1998) has been designed to handle mobility of Internet hosts at the network layer. It allows a TCP connection to remain alive and receive packets when a Mobile Host (MH) moves from one point of attachment to another Several drawbacks exist when using MIP in a mobile computing environment, the most important ones identified to date are high handover latency, high packet loss rate (Malki, 2003), and requirement for change in Internet infrastructure. Mobile IP is based on the concept of Home Agent (HA) and Foreign Agent (FA) (which requires modification to existing routers in Internet) for routing packets from previous point of attachment to the new one. An MH needs to complete the following four steps before it can receive forwarded data from the previous point of attachment: (i) perform Layer 2 (L2) handover. (ii) discover the new Care of Address (CoA), (iii) register the new CoA with the HA, and (iv) forward packets from the HA to the current CoA. During this period, the MH is unable to send or receive packets through its previous or new point of attachment (Koodli, 2004), giving rise to a large handover latency and high packet loss rate.

*The research reported in this paper was funded by NASA Grants NAG3-2528 and NAG3-2922.

MIP is known to have conflict with network security solutions (Perkins, 1998). Base MIP does not cooperate well when the HA is behind a firewall and the MH is outside the firewall, unless firewall transversal solution (Montenegro and Gupta, 1998) is used. Moreover, base MIP has difficulty in the presence of a foreign network which implements ingress filtering, unless reverse tunnelling, where the HA's IP address is used as the exit point of the tunnel, is used to send data from the MH.

1.1 Recent Research on Improving Mobile IP

Many improvements to Mobile IP have been proposed to reduce handover latency and packet loss. IP micromobility protocols like Hierarchical IP (Gustafsson et al., 2001), HAWAII (Ramjee et al., 1999) and Cellular IP (Cambell et al., 1999) use hierarchical foreign agents to reduce the frequency and latency of location updates by handling most of the handovers locally. Low latency Handoffs in Mobile IPv4 (Malki, 2003) uses pre-registrations and post-registrations which are based on utilizing link layer event triggers to reduce handover latency.

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 41–52.

Optimized smooth handoff (Perkins and Wang, 1999) not only uses a hierarchical FA structure, but also queues packets at the visited FA buffer and forward packets to MH's new location. To facilitate packet rerouting after handover and reduce packet losses, Jung et.al. (Jung et al., 2002) introduces a location database that maintains the time delay between the MH and the crossover node. Mobile Routing Table (MRT) has been introduced at the home and foreign agents in (Wu et al., 2002), and a packet forwarding scheme similar to (Perkins and Wang, 1999) is also used between FAs to reduce packet losses during handover. A reliable mobile multicast protocol (RMMP), proposed in (Liao et al., 2000), uses multicast to route data packets to adjacent subnets to ensure low packet loss rate during MH roaming. In (Fu and Atiquzzaman, 2003), Fu et al. use SCTP, a new transport layer protocol, to improve the performance of MIP by utilizing SCTP's unlimited SACK Gap Ack Blocks (Fu et al., 2005).

Mobile IPv6 (Johnson et al., 2004) removes the concept of FA to reduce the requirement on infrastructure support (only HA required). Route Optimization is built in as an integral part of Mobile IPv6 to reduce triangular routing encountered in MIPv4 (Johnson et al., 2004). Fast Handovers for Mobile IPv6 (FMIPv6) (Koodli, 2004), aims to reduce handover latency by configuring a new IP address before entering a new subnet. This results in a reduction in the time required to prepare for new data transmission; packet loss rate is thus expected to decrease. Like the Hierarchical IP in MIPv4, Hierarchical MIPv6 mobility management (HMIPv6) (Soliman et al., 2004) also introduces a hierarchy of mobile agents to reduce the registration latency and the possibility of an outdated Collocated CoA (CCOA). FMIPv6 and HMIPv6 can be used together, as suggested in (Soliman et al., 2004), to improve the performance further (in this paper, we refer to this combination as FHMIPv6). The combination of Fast Handover and HMIPv6 allows performance improvement by taking advantage of both hierarchial structure and link layer triggers. However, like FMIPv6, FHMIPv6 also relies heavily on accurate link layer information. MH's high movement speed or irregular movement pattern may reduce the performance gains of these protocols. Even with the above enhancements, Mobile IP still can not completely remove the latency resulting from the four handover steps mentioned earlier, resulting in a high packet loss rate (Hsieh and Seneviratne, 2003).

1.2 Motivation of SIGMA

As the amount of real-time traffic over wireless networks keeps growing, the deficiencies of the network layer based Mobile IP, in terms of latency and packet loss, becomes more obvious. The question that naturally arises is: Can we find an alternative approach to network layer based solution for mobility support? Since most of the applications in the Internet are end-to-end, a transport layer mobility solution would be a natural candidate for an alternative approach. A number of transport layer mobility protocols have been proposed in the context of TCP, for example, MSOCKS (Maltz and Bhagwat, 1998) and connection migration solution (Snoeren and Balakrishnan, 2000). These protocols implement mobility as an end-to-end service without the requirement to change the network layer infrastructures; they, however, do not aim to reduce the high latency and packet loss resulting from handovers. As a result, the handover latency for these schemes is in the scale of seconds.

Traditionally, various *diversity* techniques have been used extensively in wireless communications to combat channel fadings by finding independent communication paths at physical layer. Common diversity techniques include: space (or antenna) diversity, polarization diversity, frequency diversity, time diversity, and code diversity (Rappaport, 1996; Caire et al., 1998). Recently, increasing number of mobile nodes are equipped with multiple interfaces to take advantage of overlay networks (such as WLAN and GPRS) (Holzbock, 2003). The development of Software Radio technology (Glossner et al., 2003) also enables integration of multiple interfaces into a single network interface card. With the support of multiple IP addresses in one mobile host, a new form of diversity: *IP diversity* can be achieved. On the other hand, A new transport protocol proposed by IETF, called Stream Control Transmission Protocol (SCTP), has recently received much attention from the research community (Fu and Atiquzzaman, 2004). In the field of mobile and wireless communications, the performance of SCTP over wireless links (Fu et al., 2002), satellite networks (Fu et al., 2003; Atiquzzaman and Ivancic, 2003), and mobile ad-hoc networks (Ye et al., 2002) is being studied. Multihoming is a built-in feature of SCTP, which can be very useful in supporting IP diversity in mobile computing environments. Mobility protocols should be able to utilize these new hardware/software advances to improve handover performance.

The *objective* of this paper is to describe the architecture, survivability, and security of a new scheme for supporting low latency, low packet loss mobility management scheme called Transport Layer Seamless Handover (SIGMA). We also show the applicability of SIGMA to mange handoffs in space networks. Similar in principle to a number of recent transport layer handover schemes (Koh et al., 2004; Xing et al., 2002; Li, 2002), the basic idea of SIGMA is to decouple location management from data transfer, and achieve seamless handover by exploiting IP diversity to keep the old path alive during the process of setting up the new path during handover. Although

we illustrate SIGMA using SCTP, it is important to note that SIGMA can be used with other transport layer protocols that support multihoming. It can also cooperate with IPv4 or IPv6 infrastructure without any support from Mobile IP.

1.3 Contributions of Current Research

The contributions of this paper are:

- Propose and develop transport layer based seamless handover (SIGMA). Here, "seamless" means low latency and low packet loss.
- Adapt SIGMA for satellite handovers in space networks.
- Evaluate the survivability and security of SIGMA, and compare with those of MIP.

1.4 Structure of this Paper

The rest of this paper is structured as follows: First, Sec. 2 describes the basic concept of SIGMA, including handover signalling procedures, timing diagram, and location management of SIGMA. We then apply the concept of SIGMA for satellite handovers in Sec. 3. The survivability and security issues of SIGMA are evaluated in Secs. 4 and 5, respectively. Finally, concluding remarks are presented in Sec. 6.

2 ARCHITECTURE OF SIGMA

In this section, we outline SIGMA's signalling procedure for mobility management in IP networks. The procedure can be divided into five parts which will be described below. The main idea of SIGMA is to decouple location management from data transfer, and achieve seamless handover by exploiting IP diversity to keep the old path alive during the process of setting up the new path during handover.

In this paper, we illustrate SIGMA using SCTP. SCTP's multi-homing allows an association between two end points to span multiple IP addresses or network interface cards. An example of SCTP multihoming is shown in Fig. 1, where both endpoints A and B have two interfaces bound to an SCTP association. The two end points are connected through two types of links: satellite at the top and ATM at the bottom. One of the links is designated as the primary while the other can be used as a backup link in the case of failure of the primary, or when the upper layer application explicitly requests the use of the backup.

A typical mobile handover in SIGMA, using SCTP as an illustration, is shown in Fig. 2, where MH is a multi-homed node connected to two wireless access networks. Correspondent node (CN) is a node sending traffic to MH, representing services like file download or web browsing by mobile users.

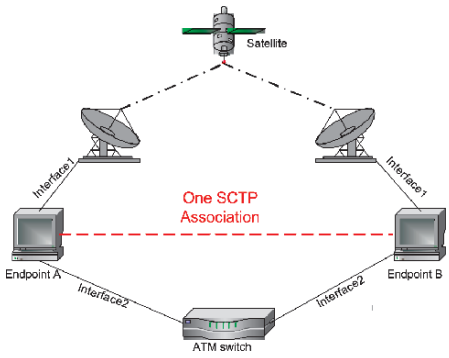


Figure 1: An SCTP association with multi-homed endpoints.

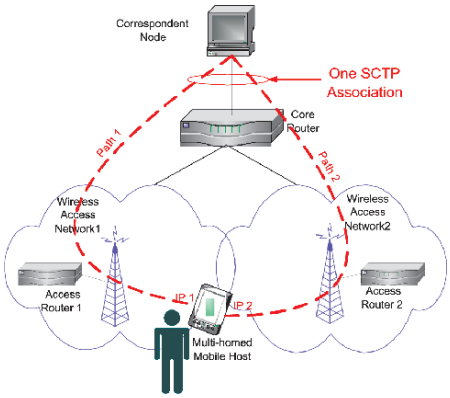


Figure 2: An SCTP association with multi-homed mobile host.

2.1 Handover Process

The handover process of SIGMA can be described by the following five steps.

STEP 1: Layer 2 handover and obtain new IP address

Refer to Fig. 2 as an example, the handover preparation procedure begins when MH moves into the overlapping radio coverage area of two adjacent subnets. In the state of the art mobile system technologies, when a mobile host changes its point of attachment to the network, it needs to perform a Layer

2 (data link layer) handover to cutoff the association with the old access point and re-associate with a new one. As an example, in IEEE802.11 WLAN infrastructure mode, this Layer 2 handover will require several steps: detection, probe, and authentication and reassociation with new AP. Only after these procedures have been finished, higher layer protocols can proceed with their signaling procedure, such as Layer 3 router advertisements. Once the MH finishes Layer 2 handover and receives the router advertisement from the new access router (AR2), it should begin to obtain a new IP address (IP2 in Fig. 2). This can be accomplished through several methods: DHCP, DHCPv6, or IPv6 stateless address auto-configuration (SAA) (Thomson and Narten, 1998).

STEP 2: Add IP addresses into the association

Initially, when the SCTP association is setup, only CN's IP address and MH's first IP address (IP1) are exchanged between CN and MH. After the MH obtained the IP address IP2 in STEP 1, MH should bind IP2 also into the association (in addition to IP1) and notify CN about the availability of the new IP address through SCTP Address Dynamic Reconfiguration option (Stewart et al., 2004). This option defines two new chunk types (ASCONF and ASCONF-ACK) and several parameter types (Add IP Address, Delete IP address, and Set Primary Address etc.).

STEP 3: Redirect data packets to new IP address

When MH moves further into the coverage area of wireless access network2, CN can redirect data traffic to new IP address IP2 to increase the possibility that data can be delivered successfully to the MH. This task can be accomplished by sending an ASCONF from MH to CN, through which CN set its primary destination address to MH's IP2. At the same time, MH need to modify its local routing table to make sure the future outgoing packets to CN using new path through AR2.

STEP 4: Update location manager (LM)

SIGMA supports location management by employing a location manager which maintains a database recording the correspondence between MH's identity and MH's current primary IP address. MH can use any unique information as its identity, such as home address (like MIP), or domain name, or a public key defined in Public Key Infrastructure (PKI).

Following our example, once MH decides to handover, it should update the LM's relevant entry with the new IP address, IP2. The purpose of this procedure is to ensure that after MH moves from wireless access network1 into network2, subsequent new association setup requests can be routed to MH's new IP address (IP2). Note that his update has no impact on the existing active associations.

We can observe an important *difference* between SIGMA and MIP: the location management and data traffic forwarding functions are coupled together in MIP, while in SIGMA they are decoupled to speedup handover and make the deployment more flexible.

STEP 5: Delete or deactivate obsolete IP address

When MH moves out of the coverage of wireless access network1, no *new* or *retransmitted* data should be directed to address IP1. In SIGMA, MH notifies CN that IP1 is out of service for data transmission by sending an ASCONF chunk to CN to delete IP1 from CN's available destination IP list.

A less aggressive way to prevent CN from sending data to IP1 is to let MH advertise a zero receiver window (corresponding to IP1) to CN. This will give CN an impression that the interface (on which IP1 is bound) buffer is full and can not receive data any more. By deactivating, instead of deleting, the IP address, SIGMA can adapt more gracefully to MH's zigzag movement patterns and reuse the previous obtained IP address (IP1) as long as the IP1's lifetime is not expired. This will reduce the latency and signalling traffic caused by obtaining a new IP address.

2.2 Timing Diagram of SIGMA

Figure. 3 summarizes the signalling sequences involved in SIGMA, the numbers before the events correspond to the step numbers in Sec. 2.1. Here we assume IPv6 SAA is used for MH to get new IP address. It should be noted that before the old IP is deleted at CN, it can receive data packets (not shown in the figure) in parallel with the exchange of signalling packets.

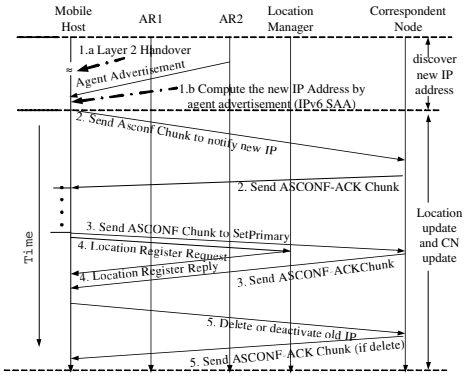


Figure 3: Timing diagram of SIGMA.

2.3 Location Management

As mentioned in STEP 4 of Sec. 2.1, SIGMA needs to setup a location manager for maintaining a database of the correspondence between MH's identity and its current primary IP address. Unlike MIP, the location

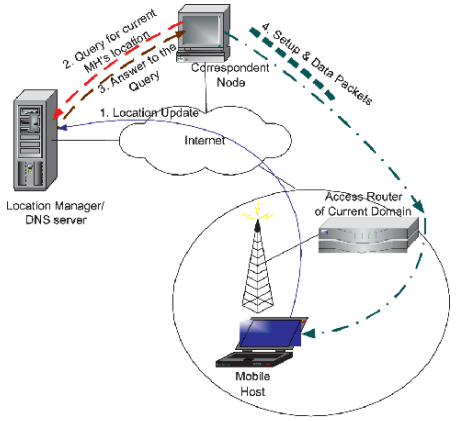


Figure 4: Location management in SIGMA.

manager in SIGMA is not restricted to the same subnet as MH's home network (in fact, SIGMA has no concept of home or foreign network). The location of the LM does not have impact on the handover performance of SIGMA. This will make the deployment of SIGMA much more flexible than MIP.

The location management can be done in the following sequence as shown in Fig. 4: (1) MH updates the location manager with the current primary IP address. (2) When CN wants to setup a new association with MH, CN sends a query to the location manager with MH's identity (home address, domain name, or public key, etc.) (3) Location manager replies to CN with the current primary IP address of MH. (4) CN sends an SCTP INIT chunk to MH's new primary IP address to setup the association.

If we use the domain name as MH's identity, we can merge the location manager into a DNS server. The idea of using a DNS server to locate mobile users can be traced back to (Awerbuch and Peleg, 1991). The advantage of this approach is its transparency to existing network applications that use domain name to IP address mapping. An Internet administrative domain can allocate one or more location servers for its registered mobile users. Compared to MIP's requirement that each subnet must have a location management entity (HA), SIGMA can reduce system complexity and operating cost significantly by not having such a requirement. Moreover, the survivability of the whole system will also be enhanced as discussed in Sec. 4.

3 SIGMA-SN: SIGMA IN SPACE NETWORKS

Spacecrafts, such as satellites, communicate among themselves and with ground stations on the earth to enable space communications. Depending on the altitude, satellites can be classified into three types: Low Earth Orbit (LEO), Medium Earth Orbit (MEO) and Geosynchronous Earth Orbit (GEO). GEO satellites are stationary with respect to earth, but LEO and MEO satellites move around the earth, and are handed over between ground stations as they pass over different areas of the earth. This is analogous to mobile computers being handed over between access points as the users move in a terrestrial network.

The National Aeronautics and Space Administration (NASA) has been studying the use of Internet protocols in spacecrafts for space communications (Bhasin and Hayden, 2002). For example, the Global Precipitation Measurement (GPM) project is studying the possible use of Internet technologies and protocols to support all aspects of data communication with spacecraft (Rash et al., 2002b). The Operating Missions as Nodes on the Internet (OMNI) (NASA, Hogie et al., 2001) project at GSFC is not only involved in prototyping, but is also testing and evaluating various IP-based approaches and solutions for space communications. Other efforts in using Internet protocols for space communications have also been reported in the literature (Minden et al., 2002).

Some of the NASA-led projects on IP in space involve handoffs in space networks. Such projects include OMNI (Hallahan, 2002; NASA,), Communication and Navigation Demonstration on Shuttle (CANDOS) mission (Hogie, 2002), and the GPM project (Rash et al., 2002a). NASA has also been working with Cisco on developing a Mobile router (Leung et al., 2001). It is also anticipated that MIP will play a major role in various space related NASA projects such as Advanced Aeronautics Transportation Technology (AATT), Weather Information Communication (WINCOMM) and Small Aircraft Transportation Systems (SATS) (Leung et al., 2001). In this section, we will investigate the use of SIGMA in space networks to support IP mobility. First, the scenarios of network layer handoffs in satellite environment is identified. Then we introduce SIGMA-SN — the mapping of SIGMA in space network.

3.1 Handoffs in a Satellite Environment

LEO satellites have some important advantages over GEO satellites for implementing IP in space. These include lower propagation delay, lower power requirements both on satellite and user terminal, more efficient spectrum allocation due to frequency reuse between satellites and spotbeams. However, due to the non-geostationary nature and fast movement of LEO satellites, the mobility management in LEO is much more challenging than in GEO or MEO.

If one of the communicating endpoint (either satellite or user terminal) changes its IP address due to the movement of satellite or mobile user, a network layer handoff is required to migrate the connection of higher level protocol (e.g. TCP, UDP, or SCTP) to the new IP address. We describe below two scenarios requiring network layer handoff in a satellite environment.

1. *Satellite as a router* (Fig. 5): When a satellite does not have any on-board equipment which generates or consumes data, but is only equipped with on-board IP routing devices, the satellite acts as a router in the Internet. Hosts are handed over from one satellite to another as the hosts come under the footprint of different satellites due to the rotation of the LEO satellites around the Earth. Referring to Fig. 5, the Fixed Host/Mobile Host (FH/MH) needs to maintain a continuous transport layer connection with the correspondent node (CN) while their attachment points change from Satellite A to satellite B. Different satellites, or even different spot-beams within a satellite, can be assigned with different IP subnet addresses. In such a case, IP address change occurs during an inter-satellite handoff, thus requiring a network layer handoff. For highly dense service areas, a spot-beam handoff may also require a network layer handoff. Previous research (Nguyen et al., 2001; Sarikaya and Tasaki, 2001) have used Mobile IPv6 to support mobility management in LEO systems, where the FH/MH and Location Manager are mapped to Mobile IP's Mobile Node and Home Agent, respectively.
2. *Satellite as a mobile host* (Fig. 6): When a satellite has on-board equipment (such as earth and space observing equipment) which generates data for transmission to workstations on Earth, or the satellite receives control signals from the control center, the satellite acts as the endpoint of the communication, as shown in Fig. 6. Although the satellite's footprint moves from ground station A to B, the satellite should maintain continuous transport layer connection with its corespondent node (CN). A network layer handoff has to be performed if the IP address of the satellite needs to be changed due to the handover between ground stations.

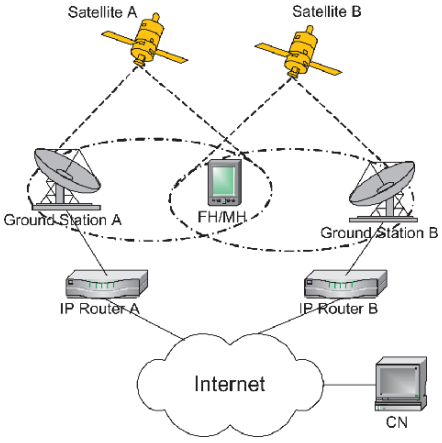


Figure 5: User handoff between satellites.

3.2 SIGMA-SN: Application of SIGMA in Space Networks

Having described our proposed SIGMA scheme and handoffs in space networks in Secs. 2 and 3.1, respectively, we describe below the mapping of SIGMA into a space handoff scenario, using satellites as examples of spacecrafts. We call this application and mapping of SIGMA to a space environment as SIGMA-SN.

1. *Satellite as a router:* Research results desribed in (Kwon and Sung, 2001) showed that the mean number of available satellites for a given FH/MH is at least two for latitudes less than 60 degrees. This means the FH/MH is within the footprint of two satellites most of the time, which makes SIGMA-SN very attractive for handoff management with a view to reducing packet loss and handoff latency. The procedure of applying SIGMA in this handoff scenario is straightforward; we just need to map the FH/MH and satellites in Fig. 5 to the MH and access routers, respectively, in the SIGMA scheme (see Fig. 2) as given below:
   - *Obtain new IP*: When FH/MH receives advertisement from Satellite B, it obtains a new IP address using either DHCP, DHCPv6, or IPv6 Stateless Address Autoconfiguration.
   - *Add new IP address to the association*: FH/MH binds the new IP address into the association (in addition to the IP address from Satellite A domain). FH/MH also notifies CN about the availability of the new IP address by sending an ASCONF chunk (Stewart et al., 2004) to the CN with the parameter type set as "Add IP Address".

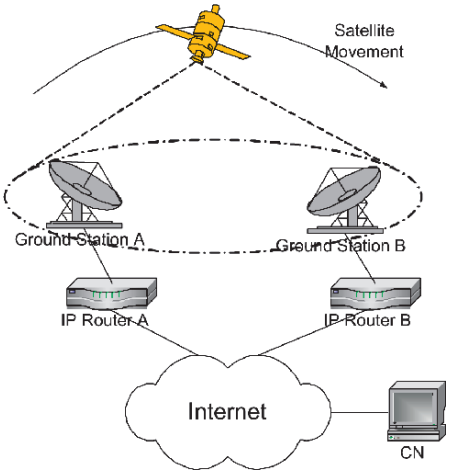


Figure 6: Satellite handoff between ground stations.

- *Redirect data packets to new IP address*: CN can redirect data traffic to the new IP address from Satellite B to increase the possibility of data being delivered successfully to the FH/MH. This task can be accomplished by sending an ASCONF chunk with the Set-Primary-Address parameter to CN, which results in CN setting its primary destination address to FH/MH as the new IP address.
- *Updating the Location manager*: `SIGMA-SN` supports location management by employing a location manager that maintains a database which records the correspondence between FH/MH's identity (such as domain name) and its current primary IP address. Once the Set-Primary-Address action is completed successfully, FH/MH updates the location manager's relevant entry with the new IP address. The purpose of this procedure is to ensure that after FH/MH moves from the footprint of Satellite A to that of Satellite B, further association setup requests can be routed to FH/MH's new IP address.
- *Delete or deactivate obsolete IP address*: When FH/MH moves out of the coverage of satellite A, FH/MH notifies CN that its IP address in Satellite A domain is no longer available for data transmission by sending an ASCONF chunk to CN with parameter type "Delete IP Address".

Due to the fixed movement track of the satellites in a space environment, FH/MH can predict the movement of Satellites A and B quite accurately. This a-priori information will be used to decide on the times to perform the set primary to the new IP address and delete the old IP address. This is much easier than in cellular networks, where the user mobility is hard to predict precisely.

2. *Satellite as a mobile host:* In this case, the satellite and IP Router A/B (see Fig. 6) will be mapped to the MH and access routers, respectively, of `SIGMA`. In order to apply `SIGMA-SN`, there is no special requirement on the Ground Stations A/B and IP routers A/B in Fig 6, which will ease the deployment of `SIGMA-SN` by not requiring any change to the current Internet infrastructure. Here, the procedure of applying `SIGMA-SN` is similar to the previous case (where the satellite acts as a router) if we replace the FH/MH by the satellite, in addition to replacing Satellites A/B by IP routers A/B.

   Since a satellite can predict its own movement track, it can contact Ground Station B while it is still connected to Ground Station A. There may be multiple new Ground Stations available to choose from due to the large footprint of satellites. The strategy for choosing a Ground Station can be influenced by several factors, such as highest signal strength, lowest traffic load, and longest remaining visibility period.

3.3 Vertical Handoff between Heterogeneous Technologies

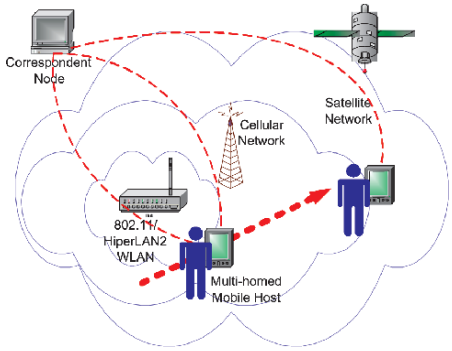


Figure 7: Vertical handoff using `SIGMA-SN`.

Different types of wireless access network technologies can be integrated to give mobile users a transparent view of the Internet. Handoff will no longer be limited to between two subnets in Wirless LAN (WLAN), or between two cells in a cellular network (horizontal handoff). In the future, mobile users will expect seamless handoff between heterogeneous

access networks (vertical handoff), such as WLANs and cellular networks.

MIP operates in Layer 3 and is independent of the underlying access network technology. Although it can be used for handoffs in a heterogeneous environment, there are a number of disadvantages in using MIP for vertical handoffs (Dixit, 2002). The disadvantages include complexity in routing, high signaling overhead, significant delay especially when CN is located in foreign network, difficulty in integrating QoS protocols such as RSVP with triangular routing and tunnelling.

SIGMA-SN is well suited to meeting the requirements of vertical handoff. Figure 7 illustrates the use of SIGMA-SN to perform vertical handoffs from WLAN to a cellular network, and then to a satellite network. A multi-homed mobile host in SIGMA-SN is equipped with multiple interface cards that can bind IP addresses obtained from different kinds of wireless network access technologies.

4 SURVIVABILITY COMPARISON OF SIGMA AND MIP

In this section we discuss the survivability of MIP and SIGMA. We highlight the disadvantages of MIP in terms of survivability, and then discuss how those issues are taken care of in SIGMA.

4.1 Survivability of MIP

In MIP, the location database of all the mobile nodes are distributed across all the HAs scattered at different locations (home networks). According to principles of distributed computing, this approach appears to have good survivability. However, there are two major drawbacks to this distributed nature of location management as given below:

- If we examine the actual distribution of the mobile users' location information in the system, we can see that each user's location and account information can only be accessible through its HA; these information are not truly distributed to increase the survivability of the system. The transparent replication of the HA, if not impossible, is not an easy task as it involves extra signaling support as proposed in (Lin and Arul, 2003).
- Even if we replicate HA to another agent, these HAs have to be located in the home network of an MH in order to intercept the packets sent to the MH. The complete home network could be located in a hostile environment, such as a battlefield, where the possibility of all HAs being destroyed is still relatively high. In the case of failure of the home networks, all the MHs belonging to the home network would no longer be accessible.

4.2 Centralized Location Management of SIGMA offers Higher Survivability

Referring to Fig. 4, SIGMA uses a centralized location management approach. As discussed in Sec. 2.1, the location management and data traffic forwarding functions in SIGMA are decoupled, allowing it to overcome many of the drawbacks of MIP in terms of survivability (see Sec. 4.1) as given below:

- The LM uses a structure which is similar to a DNS server, or can be directly combined with a DNS server. It is, therefore, easy to replicate the Location Manager of SIGMA at distributed secure locations to improve survivability.
- Only location updates/queries need to be directed to the LM. Data traffic do not need to be intercepted and forwarded by the LM to the MH. Thus, the LM does not have to be located in a specific network to intercept data packets destined to a particular MH. It is possible to avoid physically locating the LM in a hostile environment; it can be located in a secure environment, making it highly available in the network.

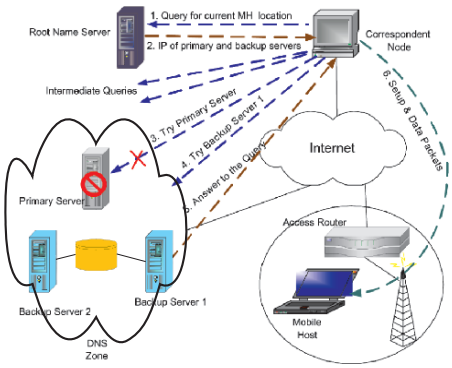


Figure 8: Survivability of SIGMA's location management.

Figure 8 illustrates the survivability of SIGMA's location management, implemented using DNS servers as location servers. Currently, there are 13 servers in the Internet (R. Bush et al., 2000) which constitute the root of the DNS name space hierarchy. There are also several delegated name servers in the DNS zone (Stevens, 1994), one of which is primary and

the others are for backup and they share a common location database. If an MH's domain name belongs to this DNS zone, the MH is managed by the name servers in that zone. When the CN wishes to establish a connection with the MH, it first sends a request to one of the root name servers, which will direct the CN to query the intermediate name servers in the hierarchy. At last, CN obtains the IP addresses of the name servers in the DNS zone to which the MH belongs. The CN then tries to contact the primary name server to obtain MH's current location. If the primary server is down, CN drops the previous request and retries backup name server 1, and so on. When a backup server replies with the MH's current location, the CN sends a connection setup message to MH. There is an important difference between the concept of MH's DNS zone in SIGMA and MH's home network in MIP. The former is a logical or soft boundary defined by domain names while the latter is a hard boundary determined by IP routing infrastructure.

If special software is installed in the primary/backup name servers to constitute a high-availability cluster, the location lookup latency can be further reduced. During normal operation, heart beat signals are exchanged within the cluster. When the primary name server goes down, a backup name server automatically takes over the IP address of the primary server. A query requests from a CN is thus transparently routed to the backup server without any need for retransmission of the request from the CN.

Other benefits SIGMA's centralized location management over MIP's location management can be summarized as follows:

- *Security*: Storing user location information in a central secure database is much more secure than being scattered over various Home Agents located at different sub-networks (in the case of Mobile IP).
- *Scalability*: Location servers do not intervene with data forwarding task, which helps in adapting to the growth in the number of mobile users gracefully.
- *Manageability*: Centralized location management provides a mechanism for an organization/service provider to control user accesses from a single server.

5 SECURITY OF SIGMA

In this section, we discuss the security issues of SIGMA and its interoperability with the current security mechanisms of the Internet.

5.1 Interoperability between MIP and Ingress Filtering

Ingress filtering is widely used in the Internet to prevent IP spoofing and Denial of Service (DoS) attacks. Ingress filtering is performed by border routers to enforce topologically correct source IP address. Topological correctness requires MH to use COA as the source IP address, since the COA is topologically consistent with the current network of the MH. On the other hand, TCP keeps track of its internal session states between communicating endpoints by using the IP address of the two endpoints and port numbers (Stevens, 1994). Therefore, applications built over TCP require the MH to always use its home address as its source address. The solution to this contradiction caused by combined requirements of user mobility, network security and transport protocols is *reverse tunnelling*, which works but lacks in terms of performance as illustrated below.

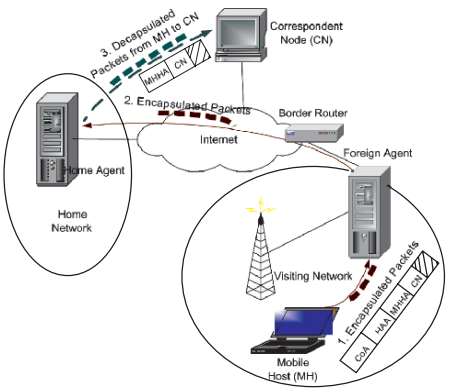


Figure 9: Interoperability between Mobile IP and Ingress Filtering.

Reverse tunnelling in MIP is shown in Fig. 9 which consists of the following components (Perkins, 2002):

1. *Encapsulation*: A data packet sent from the MH to the CN has two IP headers: the inner header has source IP address set to MH's home address (MHHA) and destination IP address set to CN's IP address; the outer header has its source IP address set to MH's CoA and destination IP address set to HA's IP address (HAA). Since the MH's CoA is topologically correct with the foreign network address, ingress filtering at foreign network's border routers allows these packets to pass through.
2. *Decapsulation*: The packets from the MH are routed towards the MH's HA because of the outer

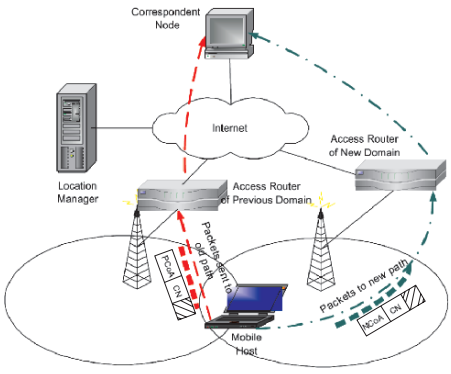


Figure 10: Interoperability between SIGMA and Ingress Filtering.

IP destination address. The HA decapsulates the packets, resulting in data packets with only one IP header (same as the previous inner header), which are then forwarded to their actual destination, i.e. the CN.

3. *Data Delivery*: When data packets arrive at the CN with the source and destination addresses being that of MH's home address and CN's address, respectively, they are identified by its TCP connection and delivered to the upper layer application.

Reverse tunnelling makes it possible for MIP to interoperate with Ingress filtering. However, the encapsulation and decapsulation of packets increase the end-to-end delay experienced by data packets, and also increase the load on the HA, which may become a performance bottleneck as the number of MHs increases.

5.2 Interoperability between SIGMA and Ingress Filtering

In SIGMA, the transport protocol uses IP diversity to handle multiple IP addresses bound to one association. The CN can thus receive IP packets from multiple source IP addresses belonging to an association, identify the association, and deliver the packets to the corresponding upper layer application. This improved capability of endpoint transport protocol permits smooth interoperability between SIGMA and Ingress Filtering.

As shown in Fig. 10, MH can use the CoA that belongs to the subnet which is responsible for sending data for the MH. In the new network, after the new CoA (NCoA) has been bound into the current association through ADDIP chunks (discussed in Sec. 2.1), the MH uses the NCoA to communicate directly with the CN. Since the NCoA is topologically correct with the subnet address, the border router of the foreign network allows packets with source IP set to the new CoA to pass. Thus, SIGMA does not require encapsulation and decapsulation as done in MIP. The transport protocol stack at the CN takes care of delivering packets coming from both previous CoA (PCoA) and NCoA to the upper layer application. SIGMA, therefore, interoperates well with ingress filtering without the need for reverse tunnelling.

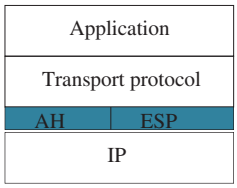


Figure 11: Use of IPSec with SCTP.

5.3 Enhancing the Security of SIGMA by IPSec

IPSec has been designed to provide an interoperable security architecture for IPv4 and IPv6. It is based on cryptography at the network layer, and provides security services at the IP layer by allowing endpoints to select the required security protocols, determine the algorithms to use, and exchange cryptographic keys required to provide the requested services. The IPSec protocol suite consists of two security protocols, namely Authentication Header (AH) and Encapsulating Security Payload (ESP). ESP provides data integrity, authentication, and secrecy services, while the AH is less complicated and thus only provides the first two services. The protocol stack, when IPSec is used with a transport protocol (SCTP in our case), is shown in Fig. 11.

SIGMA is based on dynamic address reconfiguration, which makes the association vulnerable to be hijacked, also called *traffic redirection attack*. An attacker claims that its IP address should be added into an established association between MH and CN, and further packets sent from CN should be directed to this IP address. Another kind of security risk is introduced by dynamic DNS update. An attacker can send a bogus location update to the location manager, resulting in all future association setup messages being sent to illegal IP addresses. The extra security risk introduced by SIGMA gives rise to the authentication problem: CN and LM need to determine whether the MH initiated the handover process. Since both AH and ESP support authentication, in general, we can

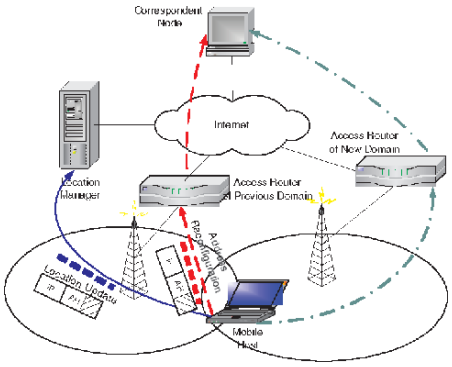


Figure 12: Interoperability between SIGMA and IPSec.

choose either of them for securing SIGMA. ESP has to be used if data confidentiality is required. Assume that we are only concerned with authentication of MH by CN and LM to prevent redirection attack and association hi-jacking. In this case, AH can be used as shown in Fig. 12. All address reconfiguration messages and location updates sent to CN and LM should be protected by IPSec AH header.

6 CONCLUSIONS

We have presented the architecture of Seamless IP diversity-based Generalized Mobility Architecture (SIGMA) to manage handovers of mobile nodes in the Internet architecture. We have shown the applicability of SIGMA to space networks for performing inter-satellite handovers, and presented the survivability and security of SIGMA. It has been shown that SIGMA has a higher survivability than MIP – thanks to its centralized location management scheme. SIGMA can also easily interoperate with existing network security infrastructures such as Ingress filtering and IPSec.

ACKNOWLEDGMENTS

We thank William Ivancic for the numerous discussion that greatly improved the quality of this paper.

REFERENCES

Atiquzzaman, M. and Ivancic, W. (2003). Evaluation of SCTP multistreaming over wireless/satellite links. In *12th International Conference on Computer Communications and Networks*, pages 591–594, Dallas, Texas.

Awerbuch, B. and Peleg, D. (1991). Concurrent online tracking of mobile users. In *ACM SIGCOMM Symposium on Communications, Architectures and Protocols*, pages 221–233.

Bhasin, K. and Hayden, J. L. (2002). Space Internet architectures and technologies for NASA enterprises. *International Journal of Satellite Communications*, 20(5):311–332.

Caire, G., Taricco, G., and Biglieri, E. (1998). Bit-interleaved coded modulation. *IEEE Transactions on Information Theory*, 44(3):927–946.

Cambell, A. T., Kim, S., and et al., J. G. (1999). Cellular IP. IETF DRAFT, draft-ietf-mobileip-cellularip-00.txt.

Dixit, S. (2002). Wireless IP and its challenges for the heterogeneous environment. *Wireless Personal Communications*, 22(2):261–273.

Fu, S. and Atiquzzaman, M. (2003). Improving end-to-end throughput of Mobile IP using SCTP. In *Workshop on High Performance Switching and Routing*, pages 171–176, Torino, Italy.

Fu, S. and Atiquzzaman, M. (2004). SCTP: State of the art in research, products, and technical challenges. *IEEE Communications Magazine*, 42(4):64–76.

Fu, S., Atiquzzaman, M., and Ivancic, W. (2002). Effect of delay spike on SCTP, TCP Reno, and Eifel in a wireless mobile environment. In *11th International Conference on Computer Communications and Networks*, pages 575–578, Miami, FL.

Fu, S., Atiquzzaman, M., and Ivancic, W. (2003). SCTP over satellite networks. In *IEEE 18th Annual Workshop on Computer Communications*, pages 112–116, Dana Point, California.

Fu, S., Atiquzzaman, M., and Ivancic, W. (2005). Evaluation of SCTP for space networks. *IEEE Wireless Communications*, 12(5):54–62.

Glossner, J., Iancu, D., Lu, J., Hokenek, E., and Moudgill, M. (2003). A software-defined communications baseband design. *IEEE Communications Magazine*, 41(1):120–128.

Gustafsson, E., Jonsson, A., and Perkins, C. (2001). Mobile IP regional registration. IETF DRAFT, draft-ietf-mobileip-reg-tunnel-04.txt.

Hallahan, F. (2002). Lessons learned from implementing Mobile IP. In *The Second Space Interent Workshop*, Greenbelt, MD.

Hogie, K. (2002). Demonstration of Internet technologies for space communication. In *The Second Space Internet Workshop*, Greenbelt, Maryland.

Hogie, K., Criscuolo, E., and Parise, R. (2001). Link and routing issues for Internet protocols in space. In *IEEE Aerospace Conference*, pages 2/963–2/976.

Holzbock, M. (2003). IP based user mobility in heterogeneous wireless satellite-terrestrial networks. *Wireless Personal Communications*, 24(2):219–232.

Hsieh, R. and Seneviratne, A. (2003). A comparison of mechanisms for improving Mobile IP handoff latency for end-to-end TCP. In *ACM MobiCom*, pages 29–41, San Diego, USA.

Johnson, D., Perkins, C., and Arkko, J. (2004). Mobility support in IPv6. IETF RFC 3775.

Jung, M., Park, J., Kim, D., Park, H., and Lee, J. (2002). Optimized handoff management method considering micro mobility in wireless access network. In *5th IEEE International Conference on High Speed Networks and Multimedia Communications*, pages 182–186.

Koh, S. J., Lee, M. J., Ma, M. L., and Tuexen, M. (2004). *Mobile SCTP for Transport Layer Mobility*. draft-sjkoh-sctp-mobility-03.txt.

Koodli, R. (2004). Fast handovers for Mobile IPv6. IETF DRAFT, draft-ietf-mipshop-fast-mipv6-03.txt.

Kwon, Y. and Sung, D. (2001). Analysis of handover characteristics in shadowed LEO satellite communication networks. *International Journal of Satellite Communications*, 19(6):581–600.

Leung, K., Shell, D., Ivancic, W., Stewart, D., Bell, T., and Kachmar, B. (2001). Application of Mobile-IP to space and aeronautical networks. *IEEE Aerospace and Electronic Systems Magazine*, 16(12):13–18.

Li, L. (2002). PKI based end-to-end mobility using SCTP. In *MobiCom 2002*, Atlanta, Georgia, USA.

Liao, W., Ke, C., and Lai, J. (2000). Reliable multicast with host mobility. In *IEEE Global Telecommunications Conference (GLOBECOM)*, pages 1692–1696.

Lin, J. and Arul, J. (2003). An efficient fault-tolerant approach for Mobile IP in wireless systems. *IEEE Transactions on Mobile Computing*, 2(3):207–220.

Malki, K. E. (2003). Low latency handoffs in Mobile IPv4. IETF DRAFT, draft-ietf-mobileip-lowlatency-handoffs-v4-07.txt.

Maltz, D. A. and Bhagwat, P. (1998). MSOCKS: An architecture for transport layer mobility. In *INFOCOM*, pages 1037–1045, San Francisco, USA.

Minden, G., Evans, J., Baliga, S., Rallapalli, S., and Searl, L. (2002). Routing in space based Internets. In *Earth Science Technology Conference*, Pasadena, CA.

Montenegro, G. and Gupta, V. (1998). Sun's SKIP firewall traversal for Mobile IP. IETF RFC 2356.

NASA. Omni: Operating missions as nodes on the internet. ipinspace.gsfc.nasa.gov.

Nguyen, H., Lepaja, S., Schuringa, J., and Vanas, H. (2001). Handover management in low earth orbit satellite IP networks. In *GlobeCom*, pages 2730–2734.

Perkins, C. (1998). Mobile Networking Through Mobile IP. *IEEE Internet Computing*, 2(1):58–69.

Perkins, C. and Wang, K. (1999). Optimized smooth handoffs in Mobile IP. In *IEEE International Symposium on Computers and Communications*, pages 340–346.

Perkins, C. E. (2002). IP Mobility Support. IETF RFC 3344.

Ramjee, R., Porta, T., and et al., S. T. (1999). IP micromobility support using HAWAII. IETF DRAFT, draft-ietf-mobileip-hawaii-00.txt.

Rappaport, T. S. (1996). *Wireless Communications Principles and Practice*. Prentice Hall, Upper Saddle River, NJ.

Rash, J., Casasanta, R., and Hogie, K. (2002a). Internet data delivery for future space missions. In *NASA Earth Science Technology Conference*, Pasadena, CA.

Rash, J., Criscuolo, E., Hogie, K., and Praise, R. (2002b). MDP: Reliable file transfer for space missions. In *NASA Earth Science Technology Conference*, Pasadena, CA.

Bush, R., Karrenberg, D., Kosters, M., and Plzak, R. (2000). Root name server operational requirements. IETF RFC 2870.

Sarikaya, B. and Tasaki, M. (2001). Supporting node mobility using mobile IPv6 in a LEO-satellite network. *International Journal of Satellite Communications*, 19(5):481–498.

Snoeren, A. C. and Balakrishnan, H. (2000). An end-to-end approach to host mobility. In *ACM MobiCom*, pages 155–166, Boston, MA.

Soliman, H., Catelluccia, C., and et al., K. M. (2004). Hierarchical Mobile IPv6 mobility management (HMIPv6). IETF DRAFT, draft-ietf-mipshop-hmipv6-04.txt.

Stevens, W. R. (1994). *TCP/IP Illustrated, Volume 1 (The Protocols)*. Addison Wesley.

Stewart, R., Ramalho, M., and et al., Q. X. (2004). Stream control transmission protocol (SCTP) dynamic address reconfiguration. IETF DRAFT, draft-ietf-tsvwg-addip-sctp-09.txt.

Thomson, S. and Narten, T. (1998). IPv6 stateless address autoconfiguration. IETF RFC 2462.

Wu, I., Chen, W., Liao, H., and Young, F. (2002). A seamless handoff approach of Mobile IP protocol for mobile wireless data networks. *IEEE Transactions on Consumer Electronics*, 48(2):335–344.

Xing, W., Karl, H., and Wolisz, A. (2002). M-SCTP: Design and prototypical implementation of an end-to-end mobility concept. In *5th Intl. Workshop on the Internet Challenge: Technology and Applications*, Berlin, Germany.

Ye, G., Saadawi, T., and Lee, M. (2002). SCTP congestion control performance in wireless multi-hop networks. In *MILCOM2002*, pages 934–939, Anaheim, California.

PART 1

Global Communication Information Systems and Services

A DECENTRALIZED LOCATION SERVICE

Applying P2P technology for picking replicas on replicated services

Luis Bernardo and Paulo Pinto
Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, P-2829-516 Caparica, Portugal
Email: lflb@uninova.pt, pfp@uninova.pt

Keywords: Scalable Internet Services and Applications, location service, self-adaptable service, peer-to-peer.

Abstract: Scalable Internet services are based on sets of peer application servers. A decentralized location service is used to resolve human readable application identifiers and return the nearest application server reference. This paper evaluates several services and algorithms from the Internet, grid and peer-to-peer community services. It identifies two potential problems and proposes a new approach for handling them. Existing techniques structure the overlay networks using tree structures. The proposed service enhances the structure with meshed structures at each level, creating dynamically multiple paths to enhance scalability. We present a study and simulation results on one aspect of scalability – sudden load of requests from users. Our service adapts to the load reaching a stable stage and performing resolution requests before a certain time limit.

1 INTRODUCTION

The growing number of users, computers, and applications servers is rapidly driving Internet to a new reality. An Internet application can no longer be a single server running on a single node. Applications must be supported by peers of machines. The convergence of web services and grid technology (Foster, 2002) provides a hint for what might be the future Internet applications.

An open architecture for applications must necessarily detach the identification of an application from the location of the application servers, in order to cope with a huge community of users and a large dynamic set of servers (peers). Such applications must not be identified by an IP address has present URLs are. Instead, an intermediate identifier must be used to bridge a human significant representation to the actual servers' location. This introduces the necessity for a middleware service that resolves the intermediate identifier to an application server reference.

This paper focuses on the implementation of such a location service. The location service must adapt to the dynamics of the application server peers and provide a scalable service to the applications. Section 2 presents an overview of the envisioned scenario. On section 3 we review several proposals of location-like services from different communities. We analyze their ability to handle simultaneous peaks of lookups and updates, and the resulting load distribution on the network. On section 4 and 5, we present our location service proposal. The paper describes its architecture and algorithms, in light of of the same requirements. We show that the introduction of a dynamic structure provides a signi ficant improvement over other approaches. The location service performance is evaluated using a set of simulations under loaded conditions on section 6.

2 PROBLEM DEFINITION

This paper assumes the existence of an active network, with a ubiquitous set of compute nodes, where application and middleware servers may run. A grid middleware layer enables the dynamic deployment of application and service servers on demand, and the access to network resources, including bandwidth and processing power. Notice though, that Internet is not homogeneous. We assume that it is composed by several interconnected high-bandwidth core networks, which interconnect a huge number of high-bandwidth and lower bandwidth networks. In order to avoid bottlenecks at the core, communication should be localized.

The location service is one of the key components for the future Internet scalability. The location service must provide an anycast resolution (Partridge, 1993) for the intermediate application identifiers. It must return the location of the application server nearest to the client. The actual

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 55–63.

metrics changes depending on the location service and on the application requirements, but it may include hops, bandwidth, stability of the nodes, processing power, etc.

The location service creates an overlay network on top of the compute nodes, which supports the application server lookup operation. The user preferred applications evolve in time. For instance, a local news service may become a top news application due to a notorious event (e.g. a local elections tie or an accident), or an e-commerce site may jump to the top due to aggressive marketing. A huge jump on the preference order may produce a huge increment on the number of clients (n_2/n_1 if Zipf distribution (Adamic, 2002) is followed). This will lead necessarily to the increase on the number of servers to cope with the demand (e.g. Content Delivery Network applications (Vakali, 2003) distribute replicas of pages to handle load peaks). A generic algorithm was proposed in (Bernardo, 1998) to control the replica deployment. The location service must be able to handle this peak of updates, and in parallel, the concurrent peak of lookups. Centralized approaches, based on a home location server may fail due to a peak of millions of requests. Caching solutions may also fail, because they may conceal the appearing of new application server replicas.

The envisioned location service provides two operations: *lookup*(*id, range*) and *update*(*id, serv_reference, range*). Each application server registers on the location service its reference associated with a unique application identifier (*id*) for a certain range. Clients search for one or more replicas within a range on the network.

3 LOCATION-LIKE SERVICES

Several existing services support the location service required functionalities. They differ on how lookups are performed: either use flooding (broadcast when available) or guided search.

Flooding approaches are common for micro-location services (e.g. Jini (Gupta, 2002)), for unstructured peer-to-peer (P2P) networks (e.g. Gnutella), and for routing algorithms in Ad Hoc networks (e.g. AODV (Perkins, 2003)). Updates are made on a local node, resulting on random information distribution. A flooding approach does not require (almost) any setting up, and adapts particularly well to unstable networks, unstable data and unstable nodes. However, it has high search costs and does not scale with the increase of the number of clients and of the lookup range (Schollmeier, 2002). Therefore, it is not adapted to provide a global view of a system. Strategies for reducing the lookup costs include (Chawathe, 2003): the creation of supernodes; the replication of information on neighbor nodes; the use of selective flooding to reduce the number of messages; and the control of the message flow. Supernodes create centralization points on a distributed network, which inter-connect lower power and more unstable nodes. They define a backbone that carries most of the flooded messages. In result, a small world effect is created that reduces the range needed to run lookups. However, supernodes also create concentration points, which can become a bottleneck on the system through link and server saturation or the increased message delay in result of flow control. Replication of *id* information distributes the load through several nodes. When replication is done at supernodes (e.g. a Clip2 Reflector replicates information for all subordinate nodes), it restricts flooding to a second hierarchical layer (connecting supernodes) with a slight increase in update costs (see table 1).

On the other hand, guided search approaches create an *id* table. Updates and lookups are made on nodes dedicated to that *id*, selected using operation *route(id)*. The table can be kept on a centralized node or partitioned and distributed on several nodes. Centralized approaches (e.g. Napster) simplify routing but introduce a single point of failure that can slow down the entire system. The performance of distributed approaches depends on the structure of *id* and on the geometry of the overlay network defined by the nodes (Gummadi, 2003). The distributed approaches include the big majority of naming and routing services and structured P2P.

DNS is a good example of the first group. DNS relies on a hierarchical structure of nodes matched with the identifier hierarchy. This approach simplifies routing because the name completely defines the resolution path. If h is the maximum hierarchical level, it has a maximum length of $2h-1$. However, it contains most of the centralized approaches limitations, benefiting only from the information fragmentation over several nodes. DNS improves its scalability using extensively caching and node replication. Caching reduces the amount of information exchanged amongst peers but prevents the use of DNS when referring to moveable or on-off entities. It was not a requirement at the time because IP addresses did not change frequently. The inflexibility of DNS routing (a single path towards the node with the required *id*) dwarfs the effects of node replication. The localization of *id* resolution (the selection of the nearest replica) is only supported by DNS extensions (e.g. Internet2 Distributed Storage Information (Beck, 1998)).

Structured P2P are based on distributed hash tables (DHT). Location servers (nodes) and

Table 1: Summary of services features: (*neighbours*) number of neighbours, (*search*) search costs, *(#path*) maximum number of independent paths available, (*update*) updates costs and (*Join*) node insertion costs. n and n_s are the average number of neighbours. N is the toal number of nodes. b and L are algorithm parameters.

	neighbours	search	# paths	update	Join
Gnutella	n	N	Path(N)	1	n
Gnut. supernode k agregation	n_s for supernodes	N/k	Path(N/k)	2	$n+1$ for nodes n_s+k for supernodes
DNS - h levels	1 node above	$\leq 2h$-1	1	$\leq 2h$-1	1 for leafs
Pastry	$O(b.\log_b N+L)$	$O(\log_b N)$	$\log_b N$	$O(\log_b N)$	$O(b.(\log_b N)^2)$
Tapestry	$O(b.\log_b N)$	$O(\log_b N)$	$\log_b N$	$O(\log_b N)$	$O(b.\log_b N)$
Brogade with k	$O(b.\log_b k+1)$ nodes	$O(\log_b (Nk))$	1 for long	$O(\log_b N)$	$O(b.\log_b k)$ for nodes

registrations are mapped to identifiers (*ids*), often calculated using hash functions. Nodes keep registrations for a subset of the *id* space. They distribute routing information creating self-organizing node structures, which exhibit some hierarchical characteristics. Each node behaves as a classical root for its local *ids*. Structured P2P services that support localization on the resolution of *id* for replicated objects include: Pastry (Rowstron, 2001), Tapestry (Hildrun, 2002), Brogade (Zhao, 2002), and other algorithms derived from Tapestry (e.g. Kademlia, AGILE).

Pastry and Tapestry are based on similar approaches. They both use strings of digits of base b as *Ids* (with a maximum N) and organize the overlay network in multiple trees (one for each *Id*). Pastry structure is a little bit more complex because it adds a complementary ring structure (L pointers), for reliability and for improving the last routing hops. However, the main routing scheme for Pastry is based on a tree. Each node is a root for the local *Ids* tree. The root connects to nodes which differ only in the last *id* digit. Successive layers differ on increasing number of digits. Nodes are members of several *id* trees in different layers. Each node maintains a routing table with $log_b\ N$ columns (one for each hierarchical level – related to the number of shared digits) and b rows (one for each digit value). Nodes route *ids* following the tree from the starting node to the root node, resulting in a maximum path length of $log_b\ N$ steps. Tapestry optimizes lookup locality for replicated sets by disseminating pointers to the application servers on the path from the node with the server till the root (associated with the *id*). Lookup is done following the direction to the *id*'s root, until a first registration is found. Pastry only replicates registrations on the direct neighbors of the root node for an *id* tree, producing longer resolution paths and less precise localization.

Brogade proposes the use of two layers of P2P overlay networks, running independent Tapestry services. Nodes with higher bandwidth connections and higher processing power are promoted to supernodes. Supernodes collect information about the lower layer node *ids* inside their region, and treat them as data on their layer. Brogade uses hierarchical routing with two levels. Users benefit from the use of more powerful links on connections. However, due to the pure hierarchy, supernodes create centralized points of failure and preclude the use of other lookup paths on the lower hierarchy level. The localization properties of Tapestry may also degrade due to the two-level routing. Local lookups use only the lower layer, but longer lookups go through the higher Tapestry network layer, resulting on a total path of $O(2log_b(k)+log_b(N/k))$.

Globe grid location service (Steen, 1998) also proposed an overlay tree structure. Globe location service trades off update flexibility for a bigger resolution path. Instead of complete information, nodes store forward pointers to other nodes (hints), which define a chain pointing to the nodes with the information. Hints usage reduces update costs because updates are propagated only to a level where another registration exists. On structured P2P updates are always propagated to the root node. Globus grid information service (Czajkowski, 2001) adopted a different approach: each node provides complete information about resources on a set of compute nodes (named a virtual organization). Nodes are organized hierarchically, feeding data upwards the hierarchy using a data access protocol and an event service. Nevertheless, if the system becomes very dynamic their choice may fail due to the update overhead.

Can the existing services support dynamic application server creation and peaks of user demand? For very dynamic information (e.g. Ad Hoc networks or fast moving objects) and for limited ranges, flooding approaches are capable of fulfilling all requirements (see update and join costs in table 1). For a scalable large scale view guided-search approaches must be adopted but they fail if the information dynamics is not slow enough. During lookup load peaks the maximum number of lookups supported depends on the number of nodes

with the information for the *id*, their capacity, but also on the total number of paths available to route to those nodes. If a concurrent update peak is also running, it also depends on update cost and on how fast updates are available to the users. Most of the structured P2P systems propose the use of information replication on neighbor nodes, and caching of previous lookup results (useless for this problem) to handle overload. However, they do not define an algorithm to dynamically perform the replication, requiring some form of external configuration. They also suffer from an intrinsic limitation of a tree organization: **it concentrates too much load on the root node**. Figure 1 illustrates the problem. If the top hierarchical layer has *k* branches, from which, *n* branches have the (application) *id* registered, then the peak lookup load on the root node is $Q(1-n\cdot k^{-1})$ for a uniformly distributed global lookup load of Q queries per second. This means that a constant fraction of the load will be handled by the root node. This paper proposes a solution to these two problems.

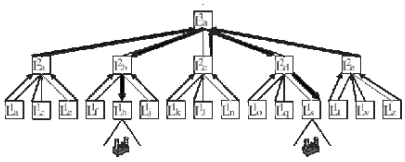


Figure 1: Load distribution on a pure tree structure.

4 LOCATION SERVICE ARCHITECTURE

This paper proposes a service location based on a dynamic tree, although enhanced with meshed structures connecting the nodes at each hierarchical level. For low load levels, the location service operates using only the paths defined by the tree. For higher load levels, extra horizontal paths are added, increasing the number of paths until the maximum supported by the overlay network. When the load is really intense new location servers can be created to split the load, and possibly, a new layer of the hierarchy can be created. This paper proposes an algorithm to control the activation of the meshed paths and the replication of nodes, based on load measurements.

The lowest layer forms a static meshed network of simple proxy entities (Local Proxy – LP) responsible to both connect to the location service and maintain neighborhood relationships. It forms a topological grid to provide a sense of "space" to the system (can be physical space, something related with availability of bandwidth between servers, etc.). Local Proxies have complete knowledge of the entities in their fixed region (be it an active node, or a set of nodes). They can resolve the identifier using the upper layers.

The layers above are composed of location servers, named L. L servers usually have hints pointing to another L server, but may have complete registration information on their first hierarchical level. L server overlay network structure is created dynamically based on the maximum range specified on *update* operations and on the load (see below). We assume that an L server can be dynamically dispatched on a particular compute node. The hierarchy is created using the clustering algorithm presented in (Bernardo, 1998b), which runs the highest hierarchical level L servers on the more resourceful compute nodes (inline with the supernode approach). As long as there are global range application servers registered, a hierarchical tree structure exists covering the entire network. Otherwise, if all applications are regional or local, there are only several independent trees. Clients can still locate *ids* using a flooding approach on the various relative roots (limited by the *range* parameter). A meshed structure connects the L servers of a tree at each hierarchical level, except the first L-server hierarchical level, which connects the entire network. But, as table 1 shows the flooding approach does not scale for a large number of clients. Therefore, we assume that application servers always specify the entire range where their clients will come from.

5 ADAPTATION OF THE RESOLUTION PATH

When an application server registers its reference, LP forwards its registration information to the first level L server. This L server disseminates a hint up the tree in direction to the root creating a single vertical path to resolve the *id*. Vertical dissemination stops when a hierarchical layer node (*h*) which embraces the required number of LPs specified in the range or when another replica hint for the same *id* is found.

A. SpreadRange

For handling the root overload problem, L servers may also disseminate hints horizontally, creating extra paths that reduce the lookup load on L servers at higher hierarchical layers. The scope of the horizontal dissemination area is defined by the

parameter SpreadRange. For instance, in fig. 2, if L^2_d SpreadRange for the illustrated application server includes L^2_e, then L^3_a stops answering to lookups coming from L^2_e (see fig. 1).

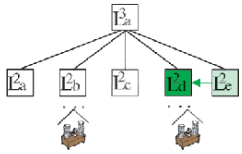


Figure 2: SpreadRange parameter.

L servers experiencing overload control horizontal dissemination for the immediately lower layer L servers, using SpreadRange Control Messages: they may send a request to some of their lower layer L servers to increment or reduce the SpreadRange on a set of identifiers. Each receiver tests its maximum range and local load, and may refuse an increase if they are higher than the maximum values allowed. When an identifier is first registered, no horizontal dissemination is used, unless a hint of an existing replica has been received from a neighbor L server. The SpreadRange parameter at an L server will be increased or decreased in result of increases and decreases of the lookup load coming from near L servers. The rationale is to deploy the structure with the lowest update overhead, yet adapted to the lookup load.

If several application server replicas are available on neighbor regions, L servers should reply to the lookups distributing the load amongst the application servers taking into account the "distance" to each of the replicas. On this case, all L servers must use the same SpreadRange value, to guarantee a balanced load distribution.

B. CoreRange

Horizontal hint dissemination can still not solve the problem because the load is still concentrated on the branch linking to the LP of the application server. Therefore, a stronger form of horizontal dissemination was introduced: the cloning of the exact L server information (and not hints) about a set of "hot" *ids* on the neighbor L servers, called replicates. The scope of replicate horizontal dissemination is controlled by the parameter CoreRange. The acceptance of a replicate is not mandatory. L servers may refuse to accept a replicate if they already are overloaded. Hence, the operation may fail. If an L server accepts the replicate, it disseminates the replicate in parallel with its local hints, vertically and possibly horizontally.

The cloning of the exact information on other L servers reduces the lookup load on the original L server. Several vertical paths can be created if CoreRange is used on several contiguous hierarchical layers. As fig. 3 shows, the lookup load is distributed between three vertical paths when CoreRange is active on the first (L^1_h) and second (L^2_b) hierarchical layers. By combining SpreadRange and CoreRange dissemination, L servers are able to control the number of paths available to resolve an *id* autonomously, creating an effective mechanism to handle peaks of load.

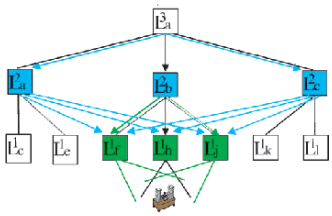


Figure 3: CoreRange parameter.

However, core meshes also increment the update overhead. In the special case of the lowest hierarchical layer, the information is the complete information and not a hint. A CoreRange modification may produce a vertical and horizontal dissemination (if SpreadRange is not null) of hints. Therefore, CoreRange will be increased or decreased in result of increases and decreases of the lookup load from far away or upper layer L servers.

When several application server replicas are available in the neighborhood, L server must distribute the load uniformly amongst them. Once more, this can only be achieved if all L servers with those hints use the same CoreRange value. L servers in the SpreadRange region forward lookups to L servers in the CoreRange region. The former servers have to know most of the latter to be able to distribute the lookups evenly. Therefore, the SpreadRanges of the L servers in the CoreRange region must cover more or less the same area. It can be proved that load can be balanced for a network with a constant number of neighbors, if L servers within SpreadRange select the forwarding L server within the CoreRange using a weight distribution algorithm, where weights decrease geometrically with the distance in a factor inversely proportional to the average number of neighbor L servers. On an unknown network it is better to use a lower factor – it is better to concentrate the load in the border of the CoreRange region than in the interior, because probably there will be more L servers at the border. With these mechanisms the original L server

receives less lookup requests from the SpreadRegion per L server, but receives from all directions.

C. L server segmentation and hierarchy

The final mechanism to handle load peaks is the creation of extra L servers at the same hierarchical layer. It reduces the number of application servers registered on some of the L servers. Additionally, if the creator L server is within the CoreRange region of another server, it increases the total processing capacity of that region, until the maximum compute power available.

If the density of L servers on a hierarchical layer becomes too high (compared to the border layers), the L server clustering algorithm may react creating new hierarchical levels. In result, the number of hierarchal levels is not uniform throughout the network: crowded areas can have deeper branches than other less crowded areas.

D. Load adaptation algorithm

An inter-L server co-ordination algorithm is used to guarantee that L servers are enough to respond to the requests and to keep the hints coherent during internal modifications. It is assumed that a percentage of the bandwidth is always available for control and signaling functions. The algorithm reacts to load measurements and registrations of identifiers using four main parameters: SpeadRange; CoreRange; the number of L servers at each hierarchical layer; and the number of hierarchical layers.

L servers monitor their local lookup load, determining if the lookups were routed from lower layer or "near" neighbors L servers (*FromDown*), or if they were routed from higher layer or "distant" L servers (*FromUp*). L servers measure their average load on fixed length intervals, using (1) (a modified discrete first order filter to attenuate the variation of the load measured). Coefficient α_i has two different values whether the load increased or decreased (α_{up} and α_{down}) compared to the previous average value. In result, the algorithm reacts more promptly to raises of the load. The load algorithm makes the system react in situations of overload (a threshold value of *MaxLoad*) and underload (*MinLoad*). L servers also monitor their queue, and react when it increases above a threshold value (*1stMaxQ*). This second mechanism speeds up response for sudden raises of load. Each time there is a reaction, a minimum interval time (*MinPeriod*) is defined to prevent a second reaction. When this time expires, the adaptation is fired again if the queue length increases above a new threshold. The new threshold takes into account the clients in the queue plus a constant increment (*MQinc*). The increase is done in such a way that the more loaded the system is (the delay increases and so does the queue) the greater is the sensitivity of the threshold (in relative terms the value has decreased) making the whole system react more often.

$$load_n = \alpha_i.measurement_n + (1-\alpha_i).load_{n-1} \quad (1)$$

When an overload trigger happens in an L server, it tries to:

```
1. if FromDown then
     Increase SpreadRange on the lower layer
     L servers;
   if FromUp then
     Increase local CoreRange;
2. If 1 failed and (load_n > NewReplicaLoad)
     Segment L server; if violates density,
     modify hierarchy
```

When an underload trigger happens in an L server, it first tries to reduce SpreadRange and CoreRange, turning the system into a more pure hierarchy. Afterwards, if $load_n$ goes below a minimum threshold, the L server tries to self-destruct. Before, it runs an agreement protocol to assess that all neighbor L servers are unloaded and select one of them to receive its lower layer L servers.

The location service parameters are also influenced by application server updates. If a server changes its location frequently the SpreadRange and CoreRange parameters will be reset frequently to zero, and in consequence, the hint dissemination is almost restricted to the vertical dimension, with low update overhead. A bigger number of application replicas produce a more uniform *id* hint distribution through the L server network, concentrating load on the lower hierarchical levels. This characteristic allows a good adaptation to extreme load peaks.

Compared to the approaches analyzed is section 3, the proposed location service presents search and update costs comparable to DNS for low load levels. When load increases, update costs are increased by a value proportional to the number of hierarchical levels (h), the extra number of paths created, and the optional horizontal load balancing costs. Notice, however, that update costs are reduced by the use of chained hints, which restrict the high priority update area to the first hierarchical L servers with complete reference information (except for deletions, which will seldom occur during overload).

6 SIMULATIONS

The location service presented on this paper was simulated using the Ptolemy simulation system (Ptolemy). The simulator implements a dynamic

application, where servers measure their request queue and react to local overload creating application server clones, in a number proportional to the ratio between the growth rate of the queue and the average service time. We assumed that application servers run in parallel without inter-synchronization. The complete application algorithm is described in (Bernardo, 1998).

Clients make a resolve request of the application *id* on the nearest L server and treat the response. The response is either a definitive one and the application server is invoked, or is a reference to another L server in which case the client repeats the process. If an L server takes more than 0.5 tics of simulation time to respond, clients go back to the previous L server. It is assumed that delays are due to load, and new application servers could appear in the meanwhile allowing the client to get a fresh one.

Application servers run one client at a time during the service time (S), and keep the remaining requests on a queue. If the application server takes more than 1.5 tics to answer to a client, the client goes back to the location service and tries to locate another application server. Again, this procedure allows clients to deal with location service adaptation, using new fresh paths that could appear.

All simulations were conducted with a network of 625 compute nodes where application and L servers run. Each node runs a LP that has an average of three connections to its neighbors, and the network has a maximum distance of 24 node hops. At time zero the location service has three hierarchical layers, with 75 L servers at the first layer, five L servers at the second layer and a single root L server at the third layer. Simulations evaluate the behavior of the location service when, at instant one (tic), a total load of 625 application clients per tic (uniformly distributed on the network) try to run the application with a single starting application server. During the simulation, L servers adapt to the load. They measure the processor utilization time during intervals of 0.5 tics and test the average load after each measurement interval (using 75% for α_{up} weight and 50% for α_{down}). *MaxLoad* and *MinLoad* thresholds were respectively 95% and 0.2%. L servers also monitor the lookup queue lengths, reacting with parameters *MinPeriod* and *MQinc*, respectively 0.2 tics and ten clients. *1stMaxQ* depended to the L server lookup service time.

In order to test scalability, the lookup service time ($1/\mu_L$) took six values raging from 1 mtic (1000 lookups per tic) to 20 mtics (50 lookups per tic). *1stMaxQ* was set to μ_L. Longer values put more stress on the location service. The largest value requires that at least thirteen L servers have references to the application server, to handle the load. Experiments with the minimum value showed that a pure hierarchy could do the job.

Three kinds of application behavior were tested: 1 - a static application server with a service time of 0.001 tics; 2 - an adaptive application behavior with service time of 0.01 tics, which originated a small set of application servers (~10); 3 - an adaptive application behavior with a service time of 0.1 tics, which originated a large set of application servers (~100). For situations 2 and 3, the time needed to create a clone of an application server was set to one tic. The scalability of the proposed algorithm can be proved by the time it takes to reduce the number of client lookups waiting on the L servers queues and the total number of clients waiting on the system (both L server queues and application server queues), respectively $tStab_L$ and *tStab*, to values lower than the rate of incoming clients (312 elements). Figure 4 shows that for every combination tested the location service stabilized. $tStab_L$ does not have a strong relation with the value of $1/\mu_L$, except for the highest values tested with a single application server. The faster reactions of the queue length triggered mechanism for loaded systems compensated the extra adaptations, showed in figure 5. On behavior 1 (single application server) a pure tree would become overloaded for values of $1/\mu_L$ above 1.6 mtics. Above this value, L servers at second and third layer became overloaded, and set SpreadRange on the first layer to its maximum allowed value (6). The value of CoreRange increased with the growth of $1/\mu_L$, reaching its maximum allowed value for $1/\mu_L = 20$ mtics (where two additional L servers were created on the first hierarchical layer). For behaviors 2 and 3 the creation of extra application servers helped to disseminate the *id* lookup load amongst several branches of the location service hierarchy, reducing the final values of CoreRange and SpreadRange.

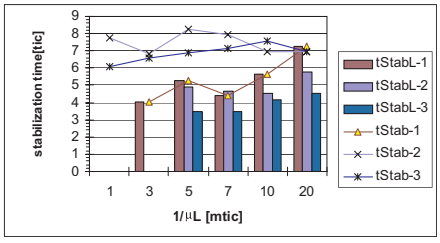


Figure 4: Stabilization times for the location service ($tStab_L$) and the total system (*tStab*).

Notice that *tStab* is almost independent of both the lookup and the application service times when

dynamic application server deployment is used. When the location service adaptation time is slower, the location service accumulates more clients on the L server queues. Once L servers react these clients are received on the application servers at a larger rate, creating a larger number of application server clones. In consequence, these clients will be processed at a higher rate.

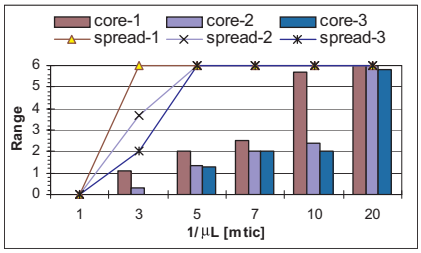


Figure 5: CoreRange and SpreadRange values at instant 100 for first layer L servers with application registrations.

7 CONCLUSIONS

Location services will play a very important role on future Internet services and applications. This paper evaluates some of the most important contributions from the P2P and grid community to solve the problem, and proposes solutions to two unhandled problems: the tree root bottleneck and the dynamic control of information replication. Previous services relied on a fixed number of paths and nodes to handle search load. Caching solved much of the overload problems however it also prevented the applications adaptation during the lifetime of cached values. The proposed service adapts to the load, creating and destroying paths on demand, in order to have the minimum update and search overhead. The update costs are reduced by the use of chained hints and by deploying just enough search paths to handle the load. Simulations results show that the solution scales with the relative increment of the load and is capable of handling concurrent search and update load.

Several other aspects of scalability could not have been addressed here: how does the entire system cope with the increase on the number of the identifiers? What are the consequences of slow and fast mobility for the coherence of the information on the servers? How large can the exchange of data be when the system is loaded due to several identifiers making the SpreadRange and the CoreRange raise toward their maximums? Another investigation subject is the support for ad hoc wireless networks. This proposal assumes that the core LP network changes seldom. For ad hoc we are investigating a new overlay structure that improves the performance of search based approaches.

REFERENCES

Adamic, A. L., Huberman, B., 2002. Zipf´s law and the Internet. In *Glottometrics* No. 3, In *http://www.ram-verlag.de*.

Beck, M., Moore, T., 1998. The Internet2 Distributed Storage Infrastructure Project: An Architecture for Internet Content Channels. Computer Networks and ISDN systems, Vol. 30, pp. 2141-2148, Nov.

Bernardo, L., Pinto, P., 1998. Scalable Service Deployment using Mobile Agents. In MA'98, the 2nd International Workshop on Mobile Agents, LNCS Vol. 1477, Springer Press.

Bernardo, L., Pinto, P., 1998b. A Scalable Location Service with Fast Update Responses. In Globecom'98, IEEE Press.

Chawathe, Y., et al., 2003. Making Gnutella-like P2P Systems Scalable. In *SIGCOMM'03, the 2003 conference on Applications, technologies, architectures, and protocols for computer communications* ACM Press.

Czajkowski, K., Fitzgerald, S., Foster, I., Kesselman, C., 2001. Grid Information Services for Distributed Resource Sharing. In HPDC'10, *the 10th IEEE International Symposium on High-Performance Distributed Computing*, IEEE Press.

Foster, I., et al., 2002. Grid Services for Distributed System Integration. IEEE Computer Vol. 35, pp. 37-46, June.

Gummadi, K., et al., 2003. The Impact of DHT Routing Geometry on Resilience and Proximity. In *SIGCOMM'03,* ACM Press.

Gupta, R., Talwar, S., Agrawal, D., 2002. Jini Home Networking: A Step toward Pervasive Computing. *IEEE Computer*, Vol. 36, pp. 34-40, Aug.

Hildrun, K., Kubiatowicz, J. D., Rao, S., Zhao, B. Y., 2002. Distributed Object Location in a Dynamic Network. In SPAA'*02, 14th annual ACM symposium on Parallel algorithms and architectures*, ACM Press.

Partridge, C., Mendez, T., Miliken, W., 1993. Host Anycasting Service. IETF RFC 1546.

Perkins, C., Belding-Royer, E., Das, S., 2003. Ad hoc On-Demand Distance Vector (AODV) Routing. IETF RFC 3561.

Ptolemy home page, http://ptolemy.eecs.berkeley.edu

Rowstron, A. I. T., Druschel, P., 2001. Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems. In Middleware'01*, 18th IFIP/ACM Int. Conf. on Distributed Systems Platforms*, LNCS Vol. 2218, Springer Press.

Schollmeier, R., Schollmeier, G., 2002. Why Peer-to-Peer (P2P) does scale: An analysis of P2P traffic patterns. In P2P'02, 2nd Int. Conf. on P2P computing, Computing, IEEE Press.

Steen, M., et al., 1998. Locating Objects in Wide-Area Systems. *IEEE Communications*, Vol. 36, pp. 104-109, Jan.

Vakali, A., Pallis, G., 2003. Content Delivery Network: Status and Trends. *IEEE Internet Computing*, Vol. 7, pp. 68-74, Nov-Dec.

Zhao, B. Y., et al., 2002. Brogade: Landmark Routing on Overlay Networks. In IPTPS'02*, the 1st Int. Workshop on Peer-to-Peer Systems*, Springer Press.

E-MACSC: A NOVEL DYNAMIC CACHE TUNING TECHNIQUE TO MAINTAIN THE HIT RATIO PRESCRIBED BY THE USER IN INTERNET APPLICATIONS

Richard S.L. Wu and Allan K.Y. Wong
Department of Computing, Hong Kong Polytechnic University, Hong Kong SAR, PRC
Email: csslwu@comp.polyu.edu.hk, csalwong@comp.polyu.edu.hk

Tharam S. Dillon
Faculty of Information Technology, University of Technology, Sydney Broadway, N.S.W. 2000
Email: tharam@it.uts.edu.au

Keywords: E-MACSC, dynamic cache tuning, popularity ratio, point-estimate, IEPM.

Abstract: The E-MACSC (*Enhanced Model for Adaptive Cache Size Control*) is a novel approach for dynamic cache tuning. The aim is to adaptively tune the cache size at runtime to maintain the prescribed hit ratio. It works with the popularity ratio (PR), defined by the standard deviations sampled for the relative popularity profile of the data objects at two successive time points. The changes in the PR value reflect the shifts of users' preference toward certain data objects. The E-MACSC makes use of the *Convergence Algorithm* (CA), which is an IEPM (*Internet End-to-End Performance Measurement*) technique that measures the mean of a waveform quickly and accurately. Accuracy of the measurement is independent/insensitive to the waveform pattern because the CA is derived from the *Central Limit Theorem*.

1 INTRODUCTION

The E-MACSC (*Enhanced Model for Adaptive Cache Size Control*) model proposed in this paper is a novel approach for dynamic cache tuning. It maintains the prescribed hit ratio for the local cache adaptively and consistently by adjusting the cache size on the fly. The adjustment is performed according to the current popularity ratio (PR). The E-MACSC is more efficacious than its MACSC (*Model for Adaptive Cache Size Control*) predecessor (Allan, 2003). The E-MACSC and MACSC tuners are especially suitable for small caching systems of limited memory resources. In fact, in the field the number of small caching systems, which usually cost less than USD$1000 (Wessels, 2001), is substantial. In these systems poor caching will lead to excess memory consumption and poor performance because of frequent task suspensions. The E-MACSC is good news for e-business applications because it shortens the RTT and keeps the customers happy. Its rationale is: "*optimal memory usage to maintain the given hit ratio*". For example, maintaining a 70% hit ratio means a RTT (*roundtrip time*) speedup of $S = RTT/(0*0.7 + RTT*0.3) \approx 3.33$. Such speedup inspires the widespread quest for different solutions to yield high hit ratios, such as the replacement strategies (Aggarwal, 1999).

2 RELATED WORK

The E-MACSC is the deeper work that bases on our previous research MACSC experience in the area of dynamic cache tuning. The focus is especially on leveraging the relative object popularity profile as the sole control metric.

2.1 Popularity Ratio

The dynamic adjustment of the cache size by the E-MACSC and MACSC models is based on the *popularity ratio* (PR) (Allan, 2003). It is the ratio of standard deviations or variances of the current popularity profile of the data objects at two consecutive time points. It is derived from the Zipf-like behaviour (Breslau, 1999; Zipf) intrinsic to cached data objects. The behaviour is represented by

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 65–72.

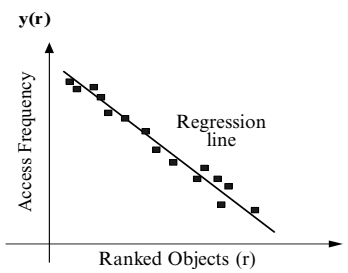


Figure 1a: Zipf-like distribution (log-log).

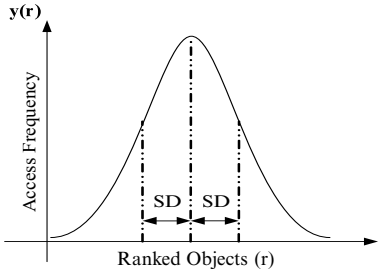


Figure 1b: Bell shape distribution. (SD – standard deviation)

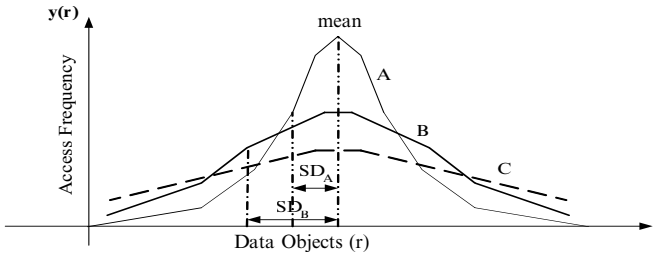


Figure 1c: Popularity distribution changes over time and reflects the change in user preference.

the log-log plot in Figure 1a, which shows that the chance (Y-axis) for the j^{th} popular object in the *sorted/ranked list* (X-axis) to be accessed is proportional to $(1/j)^{\beta}$, for $0 < \beta \leq 1$.

The original plot of the raw scattered data in Figure 1a can be approximated by the linear regression: $y(r)=f_{highest} - \gamma(r-1)$, where γ is a curve fitting parameter, r the ranked position of the object, and $f_{highest}$ the highest access frequency for the "*ranked-first*" object in the set. If this regression is mapped into the bell curve in Figure 1b, it becomes the *popularity distribution*, which shows the current profile of the relative popularity of the data objects. The central region of this curve includes the more popular objects, and $f_{highest}$ is the "*mean of the popularity distribution*" in the E-MACSC context (Allan, 2003). The shape of the popularity distribution changes over time due to the shifts in the user preference towards specific data objects. The shift is immediately reflected by the current standard deviation. For example, the three curves: A, B and C in Figure 1c mimic the different popularity distribution shapes of three different time points, and SD_A and SD_B are the standard deviations of A and B (at different time points) respectively.

2.2 The MACSC Predecessor

The MACSC tuner leverages the relative popularity of the data objects as the sole control parameter to achieve dynamic cache size tuning over the web. Leveraging the relative object popularity to heighten the hit ratio in a timely manner is a recent concept. For example, it is used as an additional parameter for the first time in the "*Popularity-Aware Greedy Dual-Size Web Proxy Caching Algorithms*" (Jin, 2000). The issue of how to utilize the relative object popularity alone for gaining higher hit ratio was never addressed before the MACSC model. As an additional parameter in a replacement algorithm (Stefan, 2003) the potential benefits from leveraging it are easily offset by the long algorithm execution time due to heavy parameterisation.

The running MACSC tuner traces all the popularity distribution changes continually and uses them timely to adjust the cache size adaptively. This tuning process maintains the prescribed minimum hit ratio consistently, and the *cache adjustment size* (CAS) is based on one of the following two equations:

$$CAS_{VR} = CacheSize_{old} * \left\{ \frac{SD_{current}}{SD_{last}} \right\}^2 \quad \text{.....(2.1)}$$

$$CAS_{SD} = CacheSize_{old} * \left(\frac{SD_{current}}{SD_{last}} \right) \quad \text{.....(2.2)}$$

The popularity ratios for equation (2.1) and equation (2.2) are the *variance ratio* (VR), and the *standard deviation ratio* (SR) (i.e. $SD_{current}$ over SD_{last}) respectively. Although the VR-based tuner (equation (2.1)) is more effective in maintaining the given hit ratio, it consumes too much memory and this makes it impractical for small caching systems with limited memory resources (Wessels, 2001; Allan, 2003). The focus here is the SR approach.

The MACSC efficacy depends on the accuracy of the popularity distribution's current standard deviation. The MACSC uses the *Point-Estimate* (PE) approach because it is derived from the Central Limit Theorem and therefore its accuracy is insensitive to traffic patterns. The $\sqrt{N}-equation$ (Chis, 1992) means the following statistical relationship:

$$E\lambda = k\delta_{\bar{x}} = k(\delta_x / \sqrt{N})$$

and the parameters are:

a) *Fractional error tolerance* (*E*): It is the error between λ (ideal/population mean value) and *m* (the mean estimated from a series of sample means $\bar{x}$ of sample size $n \geq 10$, on the fly).

b) *SD tolerance* (*k*): It is the number of standard deviations (SD) that *m* is away from the true mean λ but still be tolerated (same tolerance connotation as *E*).

c) *Predicted standard deviation* ($\delta_{\bar{x}}$): It is estimated from the same series of sample means $\bar{x}$ of sample size $n \geq 10$ by the following:

$\delta_{\bar{x}} = \delta_x / \sqrt{n}$ that fits the Central Limit Theorem.

d) *Minimum N value*: From the relationship:

$$E\lambda = k\delta_{\bar{x}} = k(\delta_x / \sqrt{N})$$

the minimum sample size *N* to compute the acceptable λ and δ_x with respect to the given *k* and *E* can be estimated. In practice $\bar{x}$ and s_x (standard deviation), estimated from the current data samples, substitute λ and δ_x as follows:

$$N = (\frac{k\delta_x}{E\lambda}) \rightarrow N = (\frac{ks_x}{E\bar{x}})^2$$

In every iteration that estimates *N* with *n* samples until $n \geq N$ convergence has occurred, the sample standard deviation s_x is estimated at the same time as $\bar{x}$. The *n* value increases with the number of estimation iterations involved till the condition: $n \geq N$ is satisfied. The PE estimates $\bar{x}$ statistically from the *n* samples first and then the standard deviation:

$$s_x = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N-1}}$$

where x_i is a data item in the i^{th} sampling round.

The following example illustrates how the PE iterative process satisfies the $n \geq N$ criterion for the $\sqrt{N}-equation$:

a) It is assumed that the initial 60 samples (i.e. sample size of n = 60) have yielded 15 and 9 for $\bar{x}$ and s_x respectively.

b) The given SD tolerance is 2 (i.e. k = 2 or 95.4%), and the fractional tolerance *E* is therefore equal to 4.6% ($E = 0.046$). Both *E* and *k* mean the same error tolerance by the *Central Limit Theorem*.

c) The minimum *N* estimate is now

$$N = (\frac{2*9}{0.046*15})^2 \approx 680$$

The value $N \approx 680$ implies that the initial sample size $n = 60$ is insufficient. To rectify the problem, one of the following methods can be adopted:

a) The first one is to collect (680 – 60) or 620 more samples and re-calculate $\bar{x}$ and s_x. There is no guarantee, however, the estimation would converge to $n \geq N$ and the same process has to be repeated.

b) The second method is to collect another 60 samples and re-calculate $\bar{x}$ and s_x from the total of 120 samples (i.e. $n = 120$ for the 2nd trial). The process repeats with 60 additional new data samples until $n \geq N$ is satisfied.

Practical experience shows that the second method converges much faster because it is common for $\bar{x}$ and s_x to stabilize in the second or third trial. This method is the basis of the core of the PE operation in the MACSC model.

The drawback of the PE process is its unpredictable time requirement to satisfy $n \geq N$ in real-life applications because of the changing IAT between any two samples. Collecting the 680 samples in the example above may take seconds, hours or even days. This kind of time un-predictability reduces the tuning precision of MACSC and makes the hit ratio oscillate in the steady state (Richard 2003). The E-MACSC model, however, does not have such unpredictability in terms of the number of data samples to satisfy $n \geq N$. Even though the IAT of the data samples can vary, the degree of severity on timeliness unpredictability is lessened because the number of data samples is fixed at *F* (the *flush limit* (equation (3.1)). This is made possible by replacing the PE process with the novel M^{3}RT *micro* IEPM or simply referred to as the

μ-IEPM mechanism. The choice of *F* is important for the fastest M³RT convergence, and the best range is: *9 ≤ F ≤ 16* (Allan, 2001). Being *micro* the mechanism operates as a logical object in the E-MACSC framework. The M³RT is a realization of the *Convergence Algorithm*, which is an IEPM technique (Cottrel, 1999) based on the *Central Limit Theorem*. For this reason the M³RT prediction accuracy is insensitive to the waveform/distribution being worked on. Previous Internet experience confirms that the M³RT mechanism always yields consistent performance even when the traffic pattern is changing continuously, for example, switching among the following patterns: Poisson, heavy-tailed, and self-similar.

2 THE E-MACSC DETAILS

The E-MACSC is an enhancement of the MACSC predecessor, which has unpredictable computation time due to: i) the unpredictable number of data samples needed by its statistical PE approach to satisfy the *N* value of the $\sqrt{N}-equation$, and ii) the unpredictable Inter-Arrival Times (IAT) among the samples (data requests). In reality, the IAT pattern affects the accuracy of the computed result because the traffic pattern can be Poisson, heavy-tailed, self-similar, or multi-fractal (Paxson, 1995).

The E-MACSC damps hit-ratio oscillation in dynamic cache tuning by replacing the PE approach with the M³RT μ-IEPM mechanism (summarized by the equations (3.1) and (3.2)), which uses *f =(F-1)* data samples to compute s_x and $\bar{x}$. The preliminary E-MACSC results indicate that the flush limit range: $9 \le F \le 16$ also yields σ_i (equation (3.6)) quickly and accurately. To enhance the sensitivity of equation (3.1) it is transformed through equation (3.3) into the equation (3.4). By arranging $\alpha = \frac{p}{p+f}$ and $(1-\alpha) = \frac{f}{p+f}$ the alternative equation (3.5) is obtained. This means that the previous $\delta_{\bar{x}}$ computation is now replaced by σ_i (equation (3.6)), and thus the PR computation is also σ_i based.

$$M_i = \frac{p * M_{i-1} + \sum_{j=1}^{j=F-1} m_i^j}{p+f} \quad \text{.....(3.1)}; i \ge 1$$

$$M_0 = m_{i=1}^{j=1} \quad \text{.....(3.2)}$$

$$\mathrm{M_i} = \frac{p}{p+f} * \mathrm{M_{i\text{-}1}} + \frac{1}{p+f}\left(\sum_{j=1}^{j=f} \mathrm{m_i^j}\right) \quad \text{.....(3.3)}$$

$$\mathrm{M_i} = \frac{p}{p+f} * \mathrm{M_{i\text{-}1}} + \frac{1}{p+f}(f)\left(\frac{1}{f}\right)\left(\sum_{j=1}^{j=f} \mathrm{m_i^j}\right) \quad \text{.....(3.4)}$$

$$\theta \quad \frac{p}{p+f} + \frac{f}{p+f} = 1$$

$$\Rightarrow \quad \mathrm{M_i} = \alpha * \mathrm{M_{i\text{-}1}} + (1-\alpha) * \left(\frac{\sum_{j=1}^{j=f} \mathrm{m_i^j}}{f}\right) \quad \text{.....(3.5)}$$

$$\sigma_i = \alpha * \sigma_{i\text{-}1} + (1-\alpha) * \sqrt{\frac{\sum_{j=1}^{j=f} \left(\mathrm{m_i^j} - \mathrm{M_i}\right)^2}{f-1}} \quad \text{..... (3.6)}$$

3 E-MACSC VERIFICATION

Many simulations were carried out with the E-MACSC prototype implemented in Java over the controlled Internet environment, as illustrated in Figure 2. The intention is to verify that the M³RT-based E-MACSC model has: a) more stable control, and b) predictably shorter execution time than its PE-based MACSC predecessor. These E-MACSC tests were carried out on the Java-based Aglets mobile agent platform [Aglets], which is chosen for its stability, rich user experience, and scalability. The Aglets platform is designed for applications over the Internet, and this makes the experimental results scalable for the open Internet. The replacement algorithm used in the simulations is the basic LRU (*Least Recently Used*) approach with the "*Twin Cache System* (TCS)" [Aggarawal99]. The TCS was used successfully in the previous MACSC verification and its function is to filter the "*one-timers*", which are considered as caching "noise". The filtration makes the hot data in the cache more concentrated by reducing the noise (Allan, 2003). One-timers are unpopular data objects that are accessed only once over a long period. The driver and the E-MACSC tuner for the proxy server in Figure 2 are aglets (*aglet applets*) that interact in a client/server relationship. The driver generates the input traffic for the E-MACSC operation, with a chosen pattern (e.g. heavy-tailed, self-similar, Poisson, and multi-fractal). The input traffic is generated by one of the following methods: a) choosing a distribution from the table (Figure 2), b) interleaving different waveforms, c) using a pre-collected data trace, or d) collecting actual data

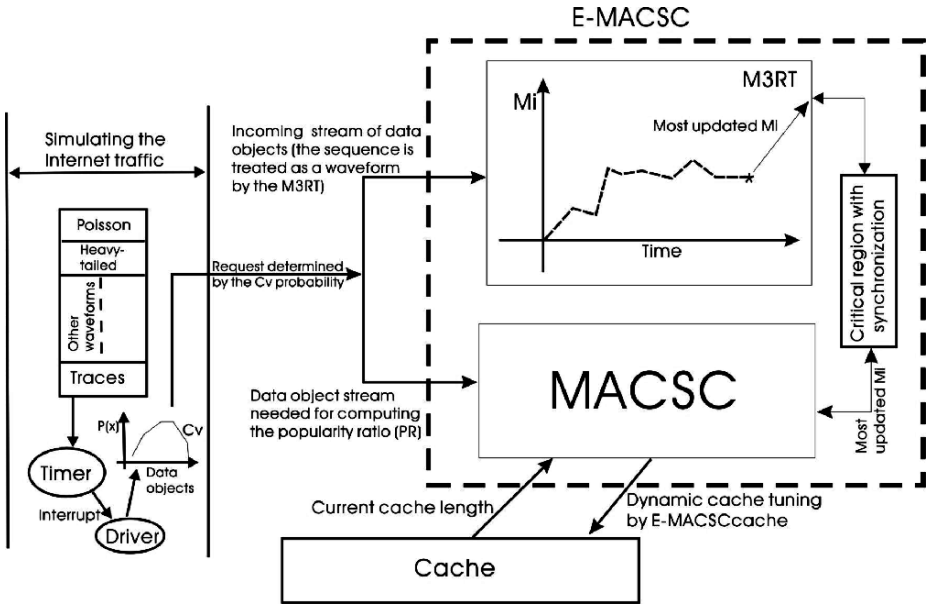


Figure 2: Verification set up for the E-MACSC tuner (for the proxy) by simulation.

object samples on the fly (Allan, 2003)). The simulation results presented in this paper are produced with the second method in two steps. The first step is to choose the waveform to drive the timer interrupt, which generates the IAT intervals. The second step is let the timer interrupt the driver so that it interpolates the unique object identifier from the X-axis of the *Cv* curve. The object identifier is then sent as the request to the server for data retrieval. The *Cv* curve is generated with either one of the four methods for the input traffic. It correlates the access probabilities/frequencies with the corresponding data objects. The object identifier is an integer (e.g. 1, 2, 3…) that uniquely identifies the specific data object.

$$CacheSize_i = CacheSize_{i-1} * \left(\frac{\sigma_i}{\sigma_{i-1}} \right) \quad \text{.....(3.7)}$$

In the verification experiments at least 40,000 data objects of various sizes are used. For example, the simulation results presented here are produced with 40,000 data objects of average size 5k bytes and 1 million data retrieval transactions generated by the driver. The cache size is first initialised to meet the prescribed hit ratio. For example, if the given hit ratio is one popularity-distribution standard deviation or 68.3%, then the initial cache size should be 5k*0.683*40,000 bytes (or 136.6MB). The basic LRU replacement algorithm deletes aged objects in the cache to accommodate the new comers. The simulation results in this paper are *SR* based (equation (2.2)), and in fact, the primary goal of the verification is to confirm that the E-MACSC tuner is more efficacious than its predecessor for small caching system applications. The preliminary results confirm that it is indeed faster and yields higher hit ratio. The popularity ratio in the E-MACSC case is computed with σ_i as shown by equation (3.7). If the given hit ratio is 68.3%, then the cache size *Z* should be initialised to $Z \approx Q * Sz$, where *Sz* is the average object size and Q equal to the number of objects that represents one standard deviation (i.e. 68.3%).

The simulation results shown in Figure 3 are produced with: 40,000 data objects of an average 5k byte size, the given hit ratio of 68.3%, 1 million data E-MACSC tuner with M^3RT support yields the highest hit ratio as compared to the "*fixed cache system* (FCS)" that works with a static cache size and the PE based MACSC tuner. Figure 4 shows the impacts by the α values, where *p* for proportional (damping) control:

$$\alpha = \frac{p}{p+f}$$

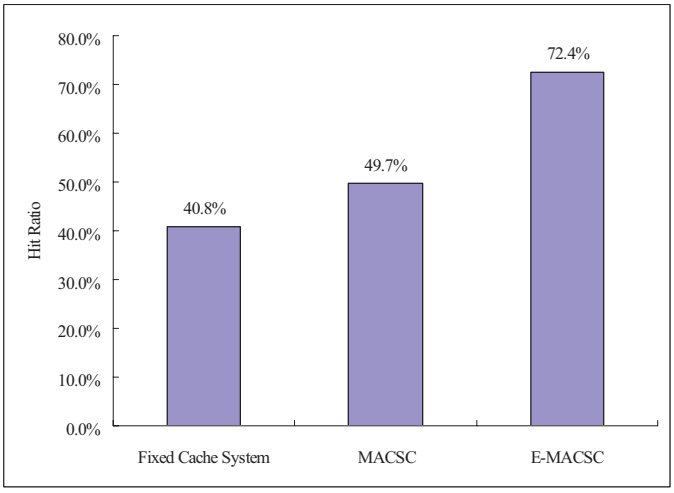


Figure 3: The comparison of the hit ratio between different algorithms. (Input traffic cyclical sequence: 3k→5k→4k→8k, α = 0.99).

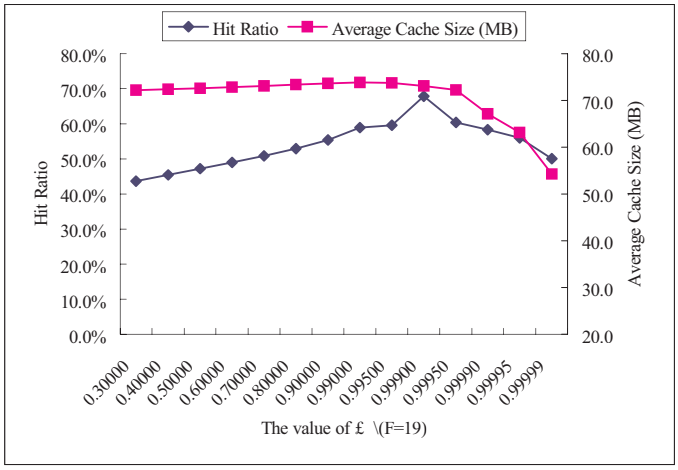


Figure 4: Hit ratio, cache size and α; F = 19, α = 0.99.

F=19 (i.e. *f=F-1=18*) produces the fastest $n \geq N$ convergence. The input traffic cyclical sequence and the given hit ratio are: *2k→6k→4k* and 68.3% (one standard deviation about the "f_{highest} *mean*" (Figure 1c)) respectively. The E-MACSC always maintains the prescribed hit ratio consistently for $a \leq 0.999$. For any α value larger than this threshold, the hit ratio drops steeply together with the memory consumption. The cause is the sudden loss of PR sensitivity because the emphasis is now on the past performance represented by α rather the current changes indicated by the *(1-α)* factor as shown in equation (3.6). Figure 5 shows the impact of different *flush limits* on the hit ratio with $a \leq 0.999$. The flush limit range that yields the highest hit ratio has shifted to $17 \leq F \leq 22$ from the best original $9 \leq F \leq 16$ range for the M_i prediction. This shift is

caused by the integrative/cumulative computation of the σ_i component in equation (3.6).

To confirm that the PE approach indeed needs more data samples to be collected on the fly to satisfy the $E\lambda = k\delta_{\bar{x}}$ criterion than the M^{3}RT mechanism, some of the above E-MACSC simulations are repeated with the MACSC under the same conditions. The average number of data samples needed by each tuner is listed in Table 1. Consistently, the MACSC tuner needs an average of 110 data samples to reach $n \geq N$ convergence, but the E-MACSC tuner needs only 18 on average. That is, the MACSC uses $(110/18) \approx 6.1$ times more samples on average and a computation overhead of 16 times, $(0.96ms/0.06ms) = 16$.

If the IAT delay and the speed of the platform are taken into account, then the average physical times to satisfy the $n \geq N$ criterion by the MACSC and E-MACSC tuners are 0.96 ms (milliseconds) and 0.06 ms respectively. The timing analysis is done with the *Intel's VTune Performance Analyzer* (VTune) and the speed of the platform being considered is 1.5 GHz (G for giga). If the average IAT is getting shorter (e.g. IAT→0), as for those simulations with pre-collected data traces where the data samples are readily usable without delay, the speedup can get up to 16 times. Yet, this is difficult to achieve in real-life applications because the data items have to be sampled one by one on the fly. It is only normal to have an IAT delay between two samples. Different simulations were carried out to verify if E-MACSC indeed is more accurate and has less oscillation than the MACSC in maintaining the given hit ratio. In the simulations the cache size under the E-MACSC control changes responsively and always settles down to satisfy the given 68.3% hit ratio requirement. For the MACSC response, however, the cache size is more oscillatory and has the tendency to stay at the higher values. There is much less chance for these problems to happen with the E-MACSC tuner because of its capability of smoother, faster, and more accurate dynamic buffer tuning.

5 CONCLUSION

The E-MACSC tuner is an enhanced successor of the previous MACSC model. It is created when the M^{3}RT μ-IEPM mechanism replaces the PE or *point-estimate* approach in the predecessor. The M^{3}RT

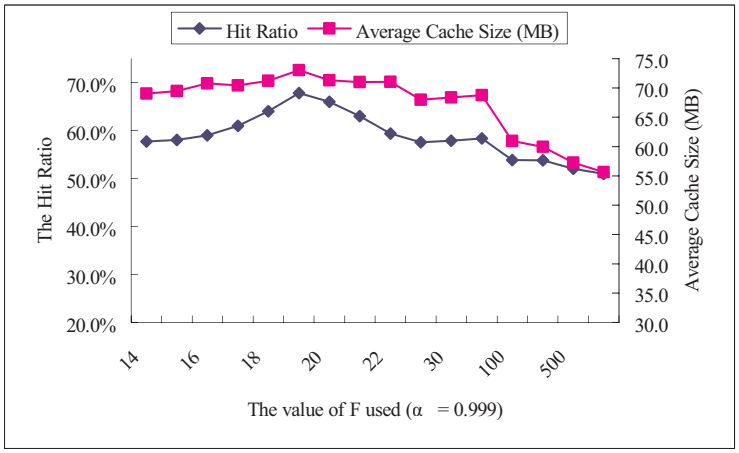


Figure 5: Correlation among hit ratio, cache size and F (α = 0.999).

Table 1: Average number of samples needed to compute the standard deviation (IAT = 0).

	Range of data samples to satisfy $n \geq N$	**Average number of data samples needed for the $n \geq N$ convergence**	**Physical time to satisfy $n \geq N$ on the platform that operates at 1.5GHz**
MACSC (PE)	60 ~ 150	110	0.96 ms
E-MACSC (M^{3}RT)	Any choice from the range: 16 ~ 20 (F value)	18 (refer to Figure 4)	0.06 ms

predicts the mean, namely, M_i of the data distribution being worked on. It differs from the PE computation by the following: a) it is faster, smoother and more accurate, b) it works with a fixed F (*flush limit*) number of data samples, and c) it is integrative with the M_{i-1} (predicted mean in the last cycle) feedback but the PE has no feedback loop. The stability of the M_i convergence, however, reduces the sensitivity of the popularity ratio that is required for producing accurate, responsive dynamic cache size tuning. With the aim to improve this sensitivity the integrative equation (3.6) for σ_i is proposed. For E-MACSC the calculation of the popularity ratio is based on equation (3.7) instead of using M_i directly. The preliminary simulation results confirm that the E-MACSC is far more efficacious in maintaining the given hit ratio than the MACSC approach. In addition the hit ratio by the E-MACSC tends to be higher than the given value. This leads to: shorter information retrieval RTT, less timeouts and thus retransmissions by the clients, more network backbone bandwidth freed for public sharing, and better system throughput in general. In contrast the hit ratio by the MACSC tuner oscillates and can be much lower than the given value. The analysis of the preliminary E-MACSC experience confirms that its efficacy depends on a few factors though, namely, the α value, the choice of the *flush limit*, and the average IAT of the data samples. Therefore, the planned activity for the next phase in the research is to study the impacts of these factors thoroughly. Different possible ways to neutralize the negative effect of some of these factors on dynamic cache tuning performance will be explored and scrutinized.

ACKNOWLEDGEMENTS

The authors thank the Hong Kong Polytechnic University and the Department of Computing for the research funding: H-JZ91

REFERENCES

Aggarwal, C., Wolf, J. L. and Philip S. Yu, 1999. Caching on the Word Wide Web. In *IEEE Transactions on Knowledge and Data Engineering*, 11(1)

Allan, K.Y. Wong, May, T.W. Ip and Richard, S. L. Wu, 2003. A Novel Dynamic Cache Size Adjustment Approach for better Retrieval Performance over the Internet, *Journal of Computer Communications*, 26(14)

Allan, K.Y. Wong, Tharam, S. Dillon, Wilfred, W.K. Lin and May, T.W. Ip, 2001. M2RT: A Tool Developed for Predicting the Mean Message Response Time for Internet Channels, *Computer Networks*, Vol. 36

Breslau, L., Cao, P., Li, F., Phillips, G. and Shenker, S., 1999. Web Caching and Zipf-like Distributions: Evidence and Implications, In *INFOCOM'99*, Vol. 1

Chis, J.A., 1992. *Introduction to Simulation and Modeling - GPSS/PC*, Prentice Hall

Cottrel, L., Zekauskas, M., Uijterwaal, H. and McGregor, T., 1999. Comparison of Some Internet Active End-to-End Performance Measurement Projects, http: //www.slac.stanford.edu/comp/net/wanmon/iepm-cf. html

Jin Shudong and Bestavros, A., 2000. Popularity-Aware Greedy Dual-Size Web Proxy Caching Algorithms, *Proc. of the Int'l Conf. on Distributed Computing Systems*

Paxson, V., Floyd, S., 1995. Wide area traffic: The Failure of Poisson Modeling, In *IEEE/ACM Transactions on Networking*, 3(3)

Richard, S.L Wu, May, T.W. Ip and Allan, K.Y. Wong, 2003. LDC-CM: A Novel Model for Dynamic Cache Size Adjustment, In *Proceeding of the International Conference on Internet Computing*, Vol. 2, Las Vegas USA

Stefan Podlipnig and Laszlo Böszörmenyi, 2003. A Survey of Web Cache Replacement Strategies, In *ACM Compputing Surveys*, Vol. 35, NO. 4, 374-398

Intel Vtune, Retrieved from http://developer.intel.com/ software/ products/vtune/

Wessels, D., 2001. Web Caching, *O'Reilly & Associates Inc.*

Zipf Curves and Website Popularity, Retrieved from http://www.useit.com/alertbox/zipf.html

EFFICIENT INFORMATION RETRIEVAL FROM HANDHELD TERMINALS WITH WIRELESS DIGITAL PHONE INTERFACE

Personalized information access on mobile phones and PDAs

Hans Weghorn
BA-University of Cooperative Education, Rotebühlplatz 41, Stuttgart, Germany
Email: weghorn@ba-stuttgart.de

Keywords: Wireless information systems, Wireless data services, Personalization, Handheld HCI, Wireless JAVA.

Abstract: Currently, the success of data services used through digital mobile phone networks is very limited. Different reasons can be identified for this: At first, the costs for data connections through these wireless networks are extremely high. Secondly, the user handling of the physically constrained handheld terminals appears as very uncomfortable. Here, a concept for customer-centred information services is proposed, which meets the limited capabilities of the terminal devices. An adequate UI is presumed to make the use of data services on mobile digital phones as also on PDAs more convenient. Furthermore, the information access speed is increased and the costs for the information retrieval are reduced by the described concept.

1 INTRODUCTION

Digital wireless telephony provides different methods for data communication. A very simple service is the exchange of short messages (SMS) as connection-less datagrams, which carry as payload a small text-based message. It is also possible to run continuous data links equivalent to a modem connection for analogue landline phone networks. For the operation of higher-level protocols, like WAP browsing, there exist today different communication methods, but the detail method is of minor interest, when using this kind of service.

What appears more important and with direct consequence to the user is how these data services are handled, and which costs have to be paid for their use. To the costs, it can be stated that these are very high in comparison to landline telephony networks. On the other hand, for an access to the worldwide Internet, private users have today mainly the choice between landline and wireless telephony networks. Most telephony companies operate both kind of networks, and the tariffs of the different companies are always in a very similar range, which ends up with the actual situation that data transfer costs through wireless links are much much more expensive than through landline networks.

Considering next, the usability of data services on handheld devices shows that the user has to deal with some inconvenient limitations. Due to their nature, the small devices are only equipped with numeric keypads and small display screens. This, of course, is required for keeping the units small, lightweight, and preserving a long operation time under battery supply. However, these inherent limitations are often disregarded by the UI structure of the terminal software: Tree-based UI systems requiring in many cases lengthy input strings are more than inappropriate for small devices. WAP browsing is one typical negative example for this: The system concept obviously was inherited from WEB browsing on desktop computers with full keyboards and huge screens, and by that, it appears not adequate for the small devices (Johnson, 1998).

In a discussion of wireless data services, WLAN as relatively new technology (IEEE standard 802.11, 1999) has to be regarded as well. For using WLAN, either a Laptop computer or at least a PDA is required, which has the proper network interface hardware. Although WLAN has established an important role in public life (Riezenman, 2002), and is available in many public places – sometimes even as cost-free service - it cannot fully substitute digital links through wireless phone networks, because in the WLAN system no roaming is foreseen. That means if the user of the wireless data link starts moving, the WLAN service is quickly lost. New WLAN standards and concepts aim to implement longer ranges and roaming features (Zahariadis, et al., 2004), but in the end, much more complicate hardware equipment is required than for the access

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 73–79.

to digital phone networks. Hence, it does not make sense to discuss WLAN at the moment as communication base for highly convenient wireless information services, which should be seamlessly accessible from any place.

Small mobile terminals like digital handheld phones or PDAs with phone interface are equipped with optimised properties in terms of movability and operation stand-by time. For instance, a typical phone can be linked to the phone network for up to two weeks without recharging its battery, and it weights only around 100 grams. On the other hand, the constraints of the user interface demand a system concept, which minimizes any required user input. This can be met by the concept of customer-centred service reported here.

In contradiction to general information services, like e.g. WAP page sets, a service has to be operated on the Internet side, which collects the information particularly desired by a certain customer. The mobile device shall play in this system only the role, which was the original intention for these units: It shall act as presentation terminal without much intelligence. Like described with the examples in the following sections there arise a series of advantages from this concept: Optimised UI handling on the terminal, reduction of data transfer volume, and by that a cost reduction of the information retrieval and an increase in information access speed.

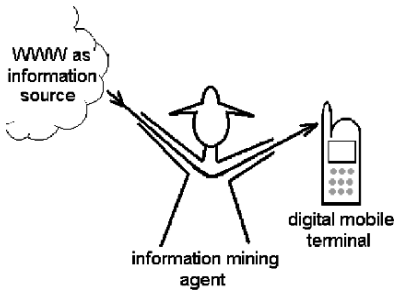


Figure 1: On demand, a specific service agent sources the desired information from the Internet and transmits the result to the mobile terminal.

2 SOLUTION APPROACH

Due to the different inherent limitations of wireless digital terminals, the following properties have to be optimised for an information retrieval operated on those units:

- Minimization of user input actions
- Minimization of data transfer in terms of volume and duration of the connection
- Minimization of computational efforts on the terminal

A system concept, which fulfils these issues, can be built up on a central service agent, which handles and prepares the desired information contents on the Internet side. In this concept, the mobile terminal has to execute simple-styled display software (Fig. 1).

2.1 Optimisation of the UI

A primary goal of the UI design is to minimize the required user actions for obtaining information contents on the mobile terminal. This can be achieved with two closely linked applications: One application is used for entering a configuration set for information queries. A second application is used for executing the information retrieval. Technically such a system can be implemented with wireless JAVA (≡ J2ME ≡ JAVA micro edition) (Piroumian, 2002). Main applications in J2ME are called MIDlets, and MIDlets can be grouped in suites for sharing a non-volatile data area – the so-called record management store. Hence, this technology is well suited for implementing exactly the proposed structure for the terminal software (Fig. 2).

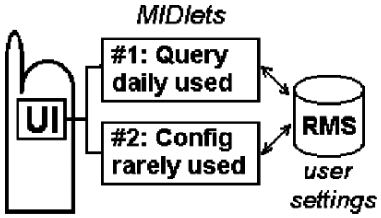


Figure 2: The terminal software can be constructed as MIDlet suite containing two independent applications: One for the information query, and another for configuring the querying parameters.

The handling properties of the two application parts are heavily asymmetric. While the actions for running the querying part are minimized, the configuration part may be as uncomfortable as used from other mobile phone software tools. For configuring the runtime behaviour, specific data has to be entered through the configuration tool. For instance,

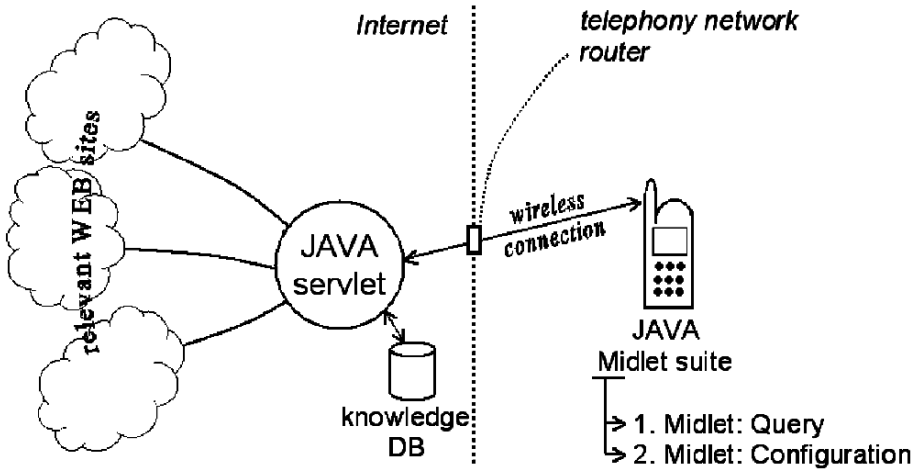


Figure 3: An intermediate information-mining agent, which can be realized as JAVA Servlet, sources the desired content from appropriate WEB sites on the Internet, and transfers the result to the handheld display terminal.

if this system should be used for accessing e-mails, all the relevant account information would have to be entered; in this sample it would be an account login and password, and a mail server address. Entering this will be very inconvenient, but the advantage is that it will be performed very seldom. On base of this input, the execution of the information queries will be very efficient and easy to use, which consequences that the user will see no barriers to use the service very often.

2.2 The Intermediate Information Service Agent

Next, the question shall be answered, how the querying tool obtains the desired information content. A data-mining agent residing on the Internet can be employed for achieving this goal (Fig. 3).

JAVA offers again a convenient technical implementation possibility with so-called Servlets (Hall, and Parr, 2001). These are small software tools, which can automatically be launched by queries to appropriately configured WEB servers. Similar technical possibilities exist alternatively with scripting programming languages like Perl or PHP, but here the question is more important, which tasks the mining agent has to fulfil.

On demand, the agent has to collect the desired information contents from the open Internet, in most cases this can come from various WEB sources. Often the same information is available on different WEB sites, which are completely independent of each other. If multi-sourcing of information is available, the central service agent can – in addition to the plain data sourcing – qualify the accuracy of the information. This quality of information – or with recent terms this would be called more QoS = quality of (information) service – can be communicated to the user. Especially the latter feature of the mining agent shows that the concept is considerably more elaborated than the idea for wrapper or mediators, which were reported earlier (Mahmoud, 2002; Wang, 2003).

Hence, the central service agent will collect the desired contents, and measure its precision, but it will also finally prepare the obtained data for a highly efficient wireless transfer. In contradiction to the coding of WEB sites with (D)HTML (Zakour, et al., 1997), or XML (Bradley, 1998), which do not regard any size limits for content pages, the wireless link should be operated with a minimal size for the transferred information. This will help to speed up the wireless transfer, and by that, it will help to minimize the network usage cost. This communication structure respects that the bottleneck in such system clearly is the wireless telephony link in terms of transfer speed and tariffs.

Figure 4: Sample of a traffic information page of an Internet WEB site of a public radio station: This source is used by the mining agent to collect the desired content - in the sample here the highway "A8" was defined as querying target. After launching the display MIDlet on the wireless terminal, the agent is activated by a HTTP query (containing as parameter the selected road), and the returned result is displayed. After this, the user can scroll through the response.

2.3 Traffic Information Channel as Practical Application Sample

For showing the practical benefit of the described concept, the following discussion should refer to an information system, which is being implemented by students of our University as exercise project. The idea of the service is to provide a personalized access to car traffic information. Most radio stations operate today WEB sites (Fig. 4). These radio stations present in their on-air program information about traffic jams, accidents, and recommended alternative driving routes. Usually, the same information is also accessible through the WEB sites, and can be used for a service according to the before described concept.

At the moment, two student teams are working in our University independently on the realization of systems, which retrieve traffic information from the WEB by a central agent, and which allow a display of this information through a specially developed software on mobile (phone) terminals. In both traffic information systems, the terminal software is implemented in JAVA according to the concept in Fig. 2. In the configuration tool the user has to enter the highway or road name, for which traffic messages shall be retrieved. For regional traffic information the QoS measure is being realized in these implementations by sourcing WEB sites of different independent radio stations, which provide this kind of messages about the same region.

A traffic query can be initiated on the mobile terminal with a minimum of user actions (Fig. 4): The query application has to be selected from the phone menu, and then it has to be launched. After this, everything happens automatically: The querying MIDlet reads the preferences from the configuration data, and activates the agent on the Internet side by a parameterised HTTP access (Knudsen, 2002). For executing the central service agent, which is realized as JAVA Servlet, a dedicate WEB server host was installed at our University with the required features. For this server a very short hostname was chosen, and a shortcut to the WEB directory link was configured, so that the services can be called by a HTTP access with a very short target address.

2.4 Realization of Other Information Systems

In the frame of an elaborated programming lecture at our University the topic of J2ME was introduced (Weghorn, 2003). The students could select as assessment work one implementation project out of a defined list of information systems, which are constructed like the one, which is described in detail above. In total twenty teams of two students were funded, and they were and they still are doing a development of information systems on car traffic channels, public transportation, railway connections, and skiing arenas in the Alps.

The student teams are implementing both parts of the information system – the terminal software, and the central service agent. Of course, since this work is part of a learning lesson, the results are not all prefect. Only a few teams came close to the above described optimal system concept. But in the end, in sum a series of information systems are implemented, which can be useful for various exemplary situations.

3 SIMULATION AND REAL WORLD DEVICES

Due to budget limitations at our University, wireless tools for digital phone networks cannot broadly be developed and tested in the target environment. This is also not really required, because simulation environments are widely available for developing and testing in particular J2ME applications. One simulator is supplied within the wireless toolkit from the company SUN Microsystems (accessible from http://java.sun.org/j2me). If the developer aims for a specific target device, all the big phone manufacturers (Motorola, Nokia, Siemens, …) operate WEB sites for developers, where simulators of various JAVA-enabled phone models can be obtained. In our practical experiments, simulation was used to develop and officially assess the many different projects.

Some of the student teams ambitiously wanted to run their tools on real physical devices and telephony networks. Hence, a few of the projects were demonstrated in the real world. During this, it turned out that the implementation of the J2ME idea is not yet sufficiently elaborated on many devices. Even when using the high-level UI only, the developer has to take care of behaviours of mobile phone devices, which were not in accordance to the original J2ME specifications. Workarounds can be used in these cases, but the problem that is shown by this experience is that at the moment it is not possible to rely on a software, which is constructed in full accordance to the J2ME specifications.

Without pointing to any specific vendor – because most phone vendors have similar styled problems – it has to be remarked that the level of the average phone J2ME has not yet reached a professional state. Furthermore, security restrictions prevent that the UI can be truly minimized, because in many phones the user has to manually select the applied network data link, when the JAVA MIDlet starts an HTTP query. By this kind of mechanism, for which there exist certainly good security arguments, the terminal software cannot be optimised to the full degree.

Another issue is the network access speed on real world devices. The query for the traffic channel information takes on a real wireless phone network approximately 15 seconds, while in the simulation environment the query is processed in around a second. Although this appears on a first glance as a back draw, the use of the described service will be much more efficient than the possible alternatives. Considering the practical example that one has to decide after breakfast which way to drive to the working office, the information collection through the proposed system would require at maximum half a minute, and around three keypad presses (≡ user input actions). The alternative would be to use a voice announcement service, which would take at least the same time, and the user has to select his particular information out of this generalised service, or to use WEB browsing on a desktop computer. Especially the latter method would consume several minutes for obtaining the desired decision input (booting of the computer, connecting to the Internet, browsing action, shutting computer down).

4 FUTURE TECHNOLOGIES

Starting the discussion of wireless data accesses with the relatively simple SMS communication system (section 1) probably does not appear as appropriate. Nevertheless, due to recent developments in J2ME technology, this ancient communication method will come into question for the described information systems. One important feature of the most recent specification of wireless JAVA (it is named MIDP version 2.0) is the capability of PUSH mechanisms (Fig. 5).

With the PUSH system, J2ME tools can be activated automatically (Ortiz, 2003). For this, the MIDlets have to register for the required service. In the particular application here, the terminal display

MIDlet can be activated by a SMS, which is sent from the Internet agent of the information system to the handheld device. Depending on the size of the transferred content, the invoking SMS can already carry the required information package, and no additional networking is required by the display MIDlet for assembling its output.

This kind of communication can be used for, e.g., the traffic information channel, because the response to the mobile terminal typically is very small: When the customer is on the way driving to the working office or back home, a fully automated system may be even more helpful than the above described one, because the user will not be able to actively handle information queries. With the PUSH system the user can be informed automatically about a recommend change of the driving route through the combination of an information agent and the appropriate terminal software. This variant would truly reach the absolute minimum for required user handling actions on the terminal.

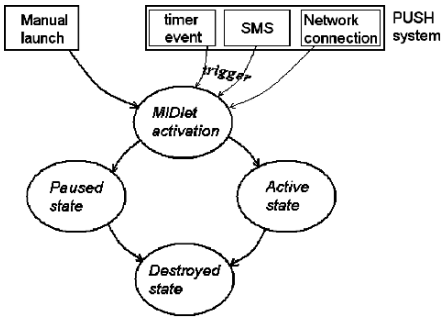


Figure 5: The life cycle of a MIDlet of the most recent J2ME generation provides with the PUSH system additional possibilities for activation.

5 CONCLUSIONS

In contradiction to generalised information services, like e.g. WAP browsing, the proposed customer-centred concept can make information access from wireless handheld terminals very comfortable and convenient. With this approach, the UI as also the locally required computational effort can be minimized. A lowering of costs and an increase of access speed is presumed to make data services used through wireless digital phone networks more attractive for the average, non-technical customer. How this can be achieved, is shown with the sample implementation of a customized traffic channel service.

The introduction of an intermediate information agent allows also accounting for a topic, which is at the moment widely disregarded: Most users are not aware of the quality of all the huge amounts of information, which can be easily retrieved by WEB browsing. The agent system allows measuring a QoS value, which is also being implemented as example in the developed traffic channel system.

As seen from the application of the proposed tools in real world networks with real phones, there is still space for optimisation. A further improvement of access speed and UI comfort can only be obtained, when the terminal developers, in particular the phone manufacturers, improve the device software. On the other hand, from the given examples it can also be derived clearly that the handling of the proposed service system is much more efficient than traditional methods through voice announcement services or WEB services accessed from desktop (or laptop) computer systems.

Further concept developments and improvements are planned with investigations on base of the newly available PUSH technologies. The goal is to further simplify the UI handling, increase the information retrieval speed, and diminish the information access costs.

REFERENCES

Bradley, N., 1998. *The XML Companion*, Addison-Wesley. Harlow, 1st edition.

Hall, M., and Parr, M., 2001. *Core Servlets and JAVA Server Pages*. Sun Microsystems Press, Prentice Hall PTR. Dorchester, 2nd edition.

IEEE standard 802.11, 1999. *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications*

Johnson, P., 1998. Usability and Mobility: Interactions on the move. In *First Workshop on Human Computer Interaction with Mobile Devices, Glasgow, May 1998*. GIST Technical Report G98-1. Department of Computing Science, University of Glasgow.

Knudsen, J., 2002. Networking, User Experience, and Threads. January 2002. *http://wireless.java.sun.com/midp/articles/threading/*

Mahmoud, Q. H., 2002. Accessing and using Internet services from JAVA-enabled handheld wireless devices. In *Braz, J., et al. (eds), 4th International Conference on Enterprise Information Systems, Ciudad Real, April 2002*. ICEIS Press.

Ortiz, E., 2003. The MIDP 2.0 Push Registry. February 2003. *http://wireless.java.sun.com/midp/articles /pushreg/*

Piroumian, V., 2002. *Wireless J2ME Platform Programming*, Prentice Hall. Palo Alto, 1st edition.

Riezenman, M. J., 2002. The ABCs of IEEE 802.11. IEEE Spectrum Online, September 2002. *http://www.spectrum.ieee.org/WEBONLY/resource/sep02/802ABCs.html*

Wang, F., et al., 2003. An E-Commerce System Integrating Data Mining Functionalities. In *Palma dos Reis, A., and Isaías, P. (eds), e-Society 2003, Lisbon, June 2003*. IADIS Press.

Weghorn, H., 2003. Projects for Lecturing Wireless JAVA. In *Palma dos Reis, A., and Isaías, P. (eds), e-Society 2003, Lisbon, June 2003*. IADIS Press.

Zahariadis, T., et al., 2004. Path Location Register for Next-Generation Heterogeneous Mobile Networks. In *Mahmoud., Q. H., and Weghorn, H. (eds), 3rd International Workshop on Wireless Information Systems WIS 2004, Porto, April 2004*. INSTICC Press.

Zakour, J., et al., 1997. *HTML 4 How-To*, Sams Publishing. Corte Madera, 1st edition.

SECURE WEB BROWSING OVER LONG-DELAY BROADBAND NETWORKS

Recommendations for Web Browsers

Doug Dillon
Assistant Vice President, Software Engineering, Hughes Network Systems
11717 Exploration Lane, Germantown, MD 20876, USA
Email:jtbs@tbs.tbs.edu

Gurjit Singh Butalia[1], Pawan Kumar Joshi[2]
[1]*Hughes Software Systems, Electronic City, Plot 31, Sector 1*
Gurgoan -122015, Haryana, India
Email: gsbutalia@hss.hns.com
[2]*Hughes Sortware Systems, 27 Gandhi Sadan*
Mandir Marg, New Delhi -110001, India
Email: pkjoshi@hss.hns.com

Keywords: Satellite Broadband, Secure Web Browsing, Performance Analysis, HTTPS, Web Browser, SSL, TLS.

Abstract: Current browser implementations provide less than desirable secure web page response time over geosynchronous satellite and other long delay broadband networks (e.g., intercontinental access across the Internet). This document defines the issues and recommends a set of enhancements that improve response time without compromising security. These enhancements are shown, by analysis, to provide more than a 50% response time reduction for a typical secure web page.

1 INTRODUCTION

Security is important on the Internet. Secure Web Browsing, in the form of the HTTP protocol running over an Secure Sockets Layer (SSL) transport (A. Frier, 1996) has proved to be the key enabling technology for E-Commerce on the Internet (Thomas, 2000) and is becoming the preferred method for secure remote access to enterprise Intranets. The SSL protocol was introduced by Netscape and has been standardized by the Internet Engineering Task Force (IETF) under the name Transport Layer Security (TLS) (Jungmaier, 2002).

The security provided by use of the HTTP protocol running over SSL transport (HTTP/SSL) comes at the cost of reduced performance. Most research and commercial product development aimed at reducing the performance impact has been focused on reducing web server processing requirements (Apostolopoulos, 2000).

The response time performance cost of HTTP/SSL has not received similar scrutiny. This paper demonstrates how current browser implementations of HTTP/SSL amplify the latency inherent in broadband networks, how this impact is felt for secure web page retrieval across networks employing either transcontinental or satellite links and provides recommendations for reducing the response time impact.

2 LONG-DELAY NETWORKS

Table 1 shows typical round-trip times over various broadband networks as measured by the authors. Consumer and enterprise satellite networks, such as the Hughes Network Systems Inc. DIRECWAY® service and the Starband consumer Internet access service, utilize demand assignment to increase the effective capacity of a satellite transponder. This introduces a second satellite round trip, which results in an overall typical round-trip time of 1300 msec.

As can be seen from the table, intercontinental Internet access round-trip time is often an order of

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 81–89.

magnitude higher than intercity, round-trip time even when no satellite links are involved.

For each of these networks, an end-user transaction that involves a single round trip provides acceptable response time for most applications.

Table 1: Measured Round-Trip Times Over Broadband Networks.

Network Type	**Ping Response Time (Sec)**
Fixed Assignment Satellite Network	.65
Demand Assigned Satellite Network	1.3
East Coast USA to India via Internet	.30
East Coast USA to Moscow via Internet	.18
Washington DC to New York	.03
Local Area Network	.001

The sections that follow analyze the number of round-trips required to retrieve a typical secure web page.

3 HTTP/SSL TRANSACTIONS

Apart from any optimizations, the HTTP or Secure Sockets Layer (HTTPS) protocol used for secure web browsing allows a web browser to retrieve a URL after four or five round-trip transactions as illustrated in figure 1. DNS lookup is required for the first retrieval of a URL from a website and, depending on where the DNS server is located within the network and the contents of its cache, may not require the response time impact of a full round trip.

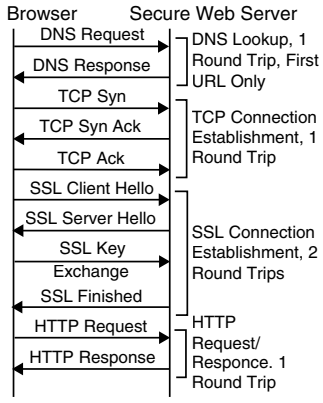


Figure 1: Unoptimized HTTP/SSL Transaction.

As illustrated by figure 2, with the SSL Session Reuse optimization (aka, session resumption), HTTPS retrieval of a URL is reduced to three round-trip transactions. A DNS lookup is not typically performed with Session Reuse as the site name was resolved when the first SSL connection to the server was established.

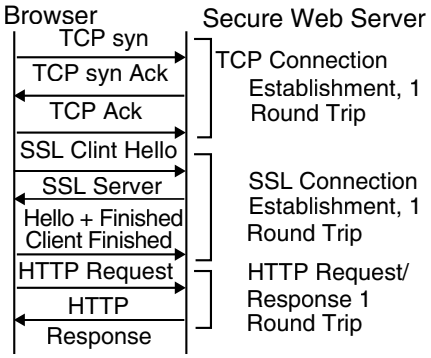


Figure 2: SSL Session Reuse.

As illustrated by figure 3, with the use of HTTP persistent connections, HTTPS retrieval of a URL is reduced to one round-trip transaction for transactions that make use of a previously SSL established connection. The use of persistent connections is not possible when the server does not support persistent connection or when the server sends back an HTTP response entity body with neither a CONTENT-LENGTH field nor chunked encoding.

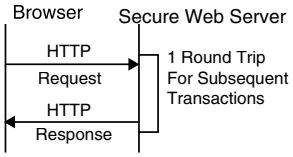


Figure 3: Persistent HTTP/SSL Connection.

4 TYPICAL SECURE WEB PAGE

Web pages become more complex as time goes by. A typical secure web page consists of multiple URLs of various kinds. This section defines a typical web page as it currently exists and then analyzes its response time. The example web page analyzed in this subsection consists of the URLs described in table 2.

Table 2: Typical Secure Web Page.

Num URLs	Description Of URL
1	HTML web page referencing two HTML frames
2	HTML page for each of the two frames
2	Cascading Style Sheet For Each HTML Frame
2	Background URL Referenced By Each Frame's Cascading Style Sheet
2	Javascript File 1 For An HTML Frame
2	Javascript File 2 For An HTML Frame
2	Redirection (301 or 302 response) to an image on another server
2	Redirected embedded images on the other server
14	Seven Embedded .gifs each HTML frame
29	**Total number of URLs**

5 PARALLELISM AND ITS LIMITS

Browsers utilize parallel HTTPS connections to retrieve URLs in parallel to reduce web page response time. For the analysis in this subsection we assume the browser is configured to support up to 16 parallel connections to each secure web server currently being accessed by the browser. Although having parallel connections make possible retrieval of URLs in parallel, other factors frequently limit the parallelism achieved. This section introduces those factors and quantifies their input.

This document refers to a "wave" of URL retrievals as a set of URLs that can be retrieved in parallel. For example, a simple web page consisting of just an HTML page and a set of embedded images may be retrieved by a browser in two waves, one for retrieving the HTML and one for retrieving the embedded images after the HTML has been parsed. This concept of a wave is an oversimplification in that it does not consider parallelism that is lost when the number of available URLs to be retrieved in parallel exceeds browser limitations on the number of parallel

Various aspects of web page design limit the achievable parallelism and increase the number of waves required to retrieve a web page. These aspects include:

- HTML Frames - the use of HTML frames limits parallelism in that a browser cannot begin to retrieve URLs referenced by an HTML frame until after that frame has been retrieved and parsed. A web page that utilizes HTML frames will always require at least three waves of URL retrievals (one for the base HTML page, one for the HTML frames, one for the items referenced by the HTML frames).
- Javascript And Cascading Style Sheets (CSS) - when a browser is parsing an HTML web page and finds a reference to a javascript URL or a Cascading Style Sheet (CSS), it typically stops parsing the HTML and retrieves the referenced URL. It resumes parsing the HTML after it has completed parsing the referenced javascript or CSS URL. This is because it is possible for the java-script or CSS file to change the way the rest of the HTML is parsed. This has the effect of limiting the parallel retrieval of URLs in that URLs referenced by the HTML cannot be retrieved until after they have been parsed and parsing is suspended while the javascript or CSS URL is being retrieved. When a browser behaves this way, a web page with N javascript and cascading stylesheet URLs may be retrieved in N + 2 waves, in the best case. This involves one wave for the HTML retrieval, one wave for each of the javascript and cascading stylesheet URLs, and one wave for embedded image retrieval.
- Javascript And CSS Referenced URLs - both Javascript and Cascading style sheets may reference URLs that are required to paint the web page. When this occurs, such a URL cannot be retrieved until after the referencing javascript or cascading style sheet has been retrieved.

Table 3: Round-Trip Time Impact on Secure Web Page Response Time.

Network Type	Ping Response Time (Sec)	Typical Secure Web Page Response Time (Sec)	Percentage of Response Time Due to Round Trips With .25 Sec Server Response Time
Fixed Assignment Satellite Network	0.650	10.4	98%
Demand Assigned Satellite Network	1.300	20.8	99%
East Coast USA to India via Internet	0.300	4.8	95%
East Coast USA to Moscow via Internet	0.180	2.88	92%
Washington DC to New York	0.030	0.48	66%
Local Area Network	0.001	0.016	6%

- Redirections - Often an embedded image or even an HTML page is actually available from a different web server than the one from which it is originally requested. A web server may instruct the browser to retrieve a URL from another web page with a redirection. A redirection may be accomplished either with an HTTP 301 or 302 response or with HTML. The use of redirection increases response time by requiring URL retrieval to obtain the redirection followed by another URL retrieval to actually obtain the URL.

6 TYPICAL WEB PAGE RESPONSE ANALYSIS

A simple table-based simulation of a web browser (see Table 5) reveals that the typical secure web page discussed is retrieved in no less than 16 round-trips even under the following set of reasonably optimistic assumptions:

• Broadband connection – i.e. the time to actually transmit packets and browser and server processing time are negligible compared to the response time contributed by the number of round trips required to paint the page.

• The browser permits up to a maximum of 16 simultaneous HTTPS connections to a given SSL server. The default value for this is for most browsers is 4 connections to HTTP 1.0 and 2 connections for HTTP 1.1. Assuming this is optimistic in that changing this configuration is for experts only in that it involves editing the windows registry, javscript files or something similarly for experts only.

• Use of persistent connections. This is enabled by default for most modern browsers and is supported in many cases by secure web servers.

• No persistent connections exist when the web page retrieval begins.

• Use of SSL session reuse. This is enabled by default for most modern browsers and servers.

• The DNS lookup for the main server does not cost a round-trip, although a DNS lookup for a redirected image's server does cost a round trip.

• Once a URL is assigned to a connection, it uses that connection even if it has to wait for the connection to be established.

• No HTTP pipelining. This paper's analysis of web page response time does not assume the use of HTTP pipelining because no popular browser has this enabled by default. This is probably due to flaws in the pipelining protocol design and server implementations thereof.

As can be seen in table 3, while round-trip time is a major contributor to response time for intercity broadband connectivity (much less so for LAN connectivity), it completely dominates inter-continental Internet and satellite network response time. This problem is significant enough for non-satellite networks that at least one Internet Startup, www.netli.com, is introducing a global distributed caching solution that advertises 1 sec secure web browsing to enterprise Intranet servers.

7 BROWSER RECOMMENDATIONS

Each of the following recommendations mitigates an inefficiency that applies especially to secure web browsing. The recommendations are as follows:

1. Connection Pooling – reduces the response time impact of SSL connection establishment whether the server supports persistent connections or not. Connection pooling establishes and maintains a pool of connections with a secure server so that an

Table 4: Improved Response Time for First and Repeat Retrievals.

Network Type	Ping Response Time (Sec)	Unoptimized Typical Secure Web Page Response Time (Sec)	Optimized Typical Web Page Response Time First Retrieval	Optimized Typical Web Page Response Time Repeat Retrieval
Fixed Assignment Satellite Network	0.650	10.4	7.80	4.55
Demand Assigned Satellite Network	1.300	20.8	15.60	9.10
East Coast USA to India via Internet	0.300	4.8	3.60	2.10
East Coast USA to Moscow via Internet	0.180	2.88	2.16	1.26
Washington DC to New York	0.030	0.48	0.36	0.21
Local Area Network	0.001	0.016	0.01	0.01
Response Time Reduction			**25%**	**56%**

established connection can be allocated to the retrieval of a URL as soon as the need to retrieve it is determined. The recommended connection pooling maintains a historical record of the number of simultaneous connections actually utilized for previous visits to a secure site and trims the connection pool size accordingly.

2. Historical Prefetch – reduces the overall response time for repeat visits to a web page by initiating the retrieval of URLs sooner than permitted by unoptimized web browsers. Historical prefetch leverages the fact that users tend to visit the same secure web pages repeatedly (for example, once a day to check a bank or brokerage account). Historical prefetch maintains a persistent cache with entries for HTML URLs. A cache entry identifies the URLs that have in the past been consistently retrieved immediately after the HTML URL was retrieved. When the retrieval of a URL in the cache commences, historical prefetch immediately initiates the retrieval of those URLs that historically were consistently retrieved immediately after the web page.

3. Quick Parse – an unoptimized browser suspends the parsing of HTML and retrieval of embedded URLs (cascading style sheets, javascript URLs, and embedded images) when a cascading stylesheet or javascript file is referenced until that URL can be retrieved and parsed. This delays the retrieval of subsequent URLs. Quick Parse quickly parses HTML to determine the set of URLs that will needed without waiting for cascading style sheets and javascript files to be loaded and parsed one at a time.

8 RECOMMENDATION RESPONSE TIME ANALYSIS

Using a tabular simulation of a browser incorporating the above recommendations (see Tables 6 and 7), allows the number of round-trips for an initial visit to the typical web page from 16 to 12. Subsequent visits to the same page require only 7 round-trips. Table 4 summarizes the impact of this reduction in round-trips for various networks.

9 CONCLUSIONS

This paper presents the reasons why current browser implementations require a large number (e.g. sixteen) network round-trips to retrieve a typical, modern secure web page. The paper provides experimental data demonstrating that these round-trip times dominate secure web page response time over transcontinental and satellite broadband networks. The paper provides recommendations to browser implementers that allow the number of round trips to be significantly reduced, especially for pages that a user repeatedly visits.

REFERENCES

Frier, A., Karton, P. and Kocher, P., 1996. *"The SSL 3.0 Protocol,"* Netscape, Nov. 1996.

Thomas, Stephen A., 2000, *"SSL & TLS Essentials: Securing the Web"*, John Wiley & Sons

A. Jungmaier, E. Rescorla, M. Tuexen., 2002, "*RFC 3436 Transport Layer Security over Stream Control Transmission"*, www.ietf.org

APPENDIX TABULAR BROWSER SIMMULATIONS

Tables 5, 6 and 7 provide the tabular simulations of browser retrieval of secure web pages without incorporating this paper's recommendations, for a first retrieval of a page when incorporating the recommendations and for a subsequent retrieval of the web page.

Table 5: Round-Trip Analysis of Typical Web Page Retrieval.

Event ID	Predecess or Event IDs	HTTPS Conn ID	Total Round Trips At End Of This Event	Event Description
1		1	1	TCP connection establishment
2	1	1	3	SSL connection establishment
3	2	1	4	Retrieval of HTML web page
4	3	1	5	Retrieval of 1st HTML frame
5	3	2	5	TCP connection establishment for 2nd HTML frame
6	4	1	6	Retrieval of 1st frame's cascading style sheet
7	5	2	6	SSL connection establishment with session reuse for 2nd HTML frame
8	6	1	7	Retrieval of 1st frame's background URL
9	7	2	7	Retrieval of 2nd HTML frame
10	6	3	7	TCP connection establishment for 1st frame's 1st javascript URL
11	8	1	8	Retrieval of 2nd frame's cascading style sheet
12	10	3	8	SSL connection establishment with session reuse for 1st frame's 1st javascript URL
13	11	1	9	Retrieval of 2nd frame's background URL
14	11	2	9	Retrieval of 2nd frame's first javascript URL
15	12	3	9	Retrieval of 1st frame's 1st javascript URL
16	13,14	1	10	Retrieval of 2nd frame's 2nd javascript URL
17	14,15	2	10	Retrieval of 1st frame's 2nd javascript URL
18	16,17	1	11	Retrieval of 1st frame's redirection to an image on another server
19	16,17	2	11	Retrieval of 2nd frame's redirection to an image on another server
20	15,17	3	11	Retrieval of 1st frame's 1st embedded image
21	17	4..9	11	TCP connection establishment 1st frame's 2nd through 7th images
22	16	10..16	11	TCP connection establishment for the 2nd frame's seven embedded images
23	21	4..9	12	SSL connection establishment with session reuse for 1st frame's 2nd through 7th images

24	22	10..16	12	SSL connection establishment with session reuse for 2nd frame's seven embedded images
25	18	17	12	DNS lookup to other server for 1st frame's redirected image
26	19	18	12	DNS lookup to other server for 2nd frame's redirected image
27	23	4..9	13	Retrieval of 2nd through 7th embedded images
28	24	10..16	13	Retrieval of 2nd frame's seven embedded images
29	25	17	13	TCP connection establishment to other server for 1st frame's redirected image
30	26	18	13	TCP connection establishment to other server for 2nd frame's redirected image
31	29	17	15	SSL connection establishment without session reuse to the other server for retrieval of the 1st frame's redirected image
32	30	18	15	SSL connection establishment without session reuse to the other server for retrieval of the 1st frame's redirected image
33	31	17	16	Retrieval of 1st frame's redirected image
34	32	18	16	Retrieval of 2nd frame's redirected image
			16	**Grand total of 16 round trips to paint page**

Table 6: Optimized First Retrieval Web Page Response Time.

Event ID	**Predecessor Event IDs**	**HTTPS Conn ID**	**Total Round Trips at End of this Event**	**Event Description**
1		1	1	TCP connection establishment
2		1	3	SSL connection establishment
3	2	2..4	4	Pooled TCP connection establishment
4	2	1	4	Retrieval of HTML web page
5	4	1	5	Retrieval of 1st HTML frame
6	3	2..4	5	Pooled SSL connection establishment with session reuse
7	4	5..16	5	Pooled TCP connection establishment triggered by need for more than one connection
8	6	1	6	Retrieval of 1st frame's cascading style sheet
9	6, 4	2	6	Retrieval of 2nd HTML frame
10	6, 4	3..4	6	Retrieval of 1st frame's 1st and 2nd javascript URLs
11	7	5..16	6	Pooled SSL connection establishment with session reuse
12	8, 9	1	7	Retrieval of 2nd frame's cascading style sheet
13	9, 10	2..3	7	Retrieval of 2nd frame's 1st and 2nd javascript URLs
14	10, 5	4	7	Retrieval of 1st frame's redirection to image on other server
15	11, 9	5	7	Retrieval of 2nd frame's redirection to image on other server

16	11, 5	6..12	7	Retrieval of 1st frame's embedded images
17	11, 9	13..16	7	Retrieval on 2nd frame's 1st through 4th embedded images
18	12, 13, 5	1..3	8	Retrieval of 2nd frame's 5th through 7th embedded images
19	14, 8	4	8	Retrieval of 1st frame's background URL
20	14	17	8	DNS lookup for retrieval of 1st frame's redirected image
21	15	18	8	DNS lookup for retrieval of 2nd frame's redirected image
22	20	17	9	TCP connection establishment for retrieval of 1st frame's redirected image.
23	21	18	9	TCP connection establishment to other server to retrieve 2nd frame's redirected image
24	22	17	11	SSL connection establishment for retrieval of 1st frame's redirected image
25	23	18	11	SSL connection establishment to other server with session reuse to retrieve 2nd frame's redirected image
26	24	17	12	Retrieval of 1st frame's redirected image
27	25	18	12	Retrieval of 2nd frame's redirected image
			12	**Grand total of 12 round trips to paint page**

Table 7: Optimized Web Page Repeat Retrieval

Event ID	**Predecess or Event IDs**	**HTTP S Conn ID**	**Total Round Trips at End of this Event**	**Event Description**
1		1	1	TCP connection establishment
2		2	1	DNS lookup to retrieve 1st frame's redirected image on other server
3		3	1	DNS lookup to retrieve 2nd frame's redirected image on other server
4	2	2	2	TCP connection establishment retrieve to 1st frame's redirected image on other server
5	3	3	2	TCP connection establishment retrieve to 1st frame's redirected image on other server
6	1	1	3	SSL connection establishment
7	6	1	4	Retrieval of HTML web page
8	4	2	4	SSL connections establishment to retrieve 1st frame's redirected image on other server.
9	5	3	4	SSL connections establishment to retrieve 1st frame's redirected image on other server.
10	2	4..18	4	Pooled TCP connection establishment triggered by need for more than one connection which is triggered by historical prefetches being queued up
11	7	1	5	Retrieval of 1st HTML frame

12	8	2	5	Retrieval of 1st frame's redirected image
13	9	3	5	Retrieval of 2nd frame's redirected image
14	10	4..18	5	Pooled SSL connection establishment with session reuse
15	11, 14	1, 4..7	6	Retrieval of 1st frame's cascading style sheet, 1st javascript, 2nd javascript, background URL and redirection to image on another server
16	14	8..13	6	Retrieval of 2nd HTML frame, 2nd frame's cascading style sheet, 1st javascript, 2nd javascript, background URL and redirection to image on another server
17	14	14..18	6	Retrieval of 1st frame's 1st through 5th embedded images
18	15	1, 4	7	Retrieval of 1st frame's 6th and 7th embedded images
19	15, 16	5..11	7	Retrieval of 2nd frame's seven embedded images
			7	**Grand total of 7 round trips to paint page**

EXPERIMENTAL BASED TOOL CALIBRATION USED FOR ASSESSING THE QUALITY OF E-COMMERCE SYSTEMS

Antonia Stefani[1], Dimitris Stavrinoudis[1], Michalis Xenos[1,2]
[1] *School of Sciences & Technology,Hellenic Open University,23 Sachtouri Str, Patras, Greece*
Email: stefani@eap.gr, stavrino@eap.gr, xenos@eap.gr
[2] *Research Academic Computer Technology Institute, 61 Riga Feraiou Str, Patras, Greece*
Email: xenos@cti.gr

Keywords: E-commerce Systems, Measurements, Quality Assessment.

Abstract: This paper presents a method used to evaluate the quality of e-commerce systems. The presented method uses a Belief Network in order to model the factors and criteria affecting the quality of e-commerce systems. This model can be applied not only for assessing the quality of e-commerce systems, but also for ensuring quality design before development. It also offers numerical results for the overall quality of an e-commerce system, as well as for its intermediate factors and lower-lever criteria. This paper presents the experimental results and the data analysis that aided towards the calibration of the model, i.e. assessing an e-commerce system and its individual characteristics based on the numerical results derived from the model.

1 INTRODUCTION

E-commerce systems have been developed at a staggering rate in recent years. In particular, they offer a full range of functions and services in order to fulfill the end-users requirements and to provide them high service quality. However, the quality of e-commerce systems is strongly related to the quality of the interface, as it is perceived by the end-user, who is also the e-customer of the system.

E-commerce system research has examined different issues of interface design, especially in Business to Consumer (B2C) systems. Emphasis was placed on usability issues (Nielsen, 2000), interface design principles (Lohse, 1998; Schafer, 2001) and end-users' behavioral model (Wilson, 2003; Sherman, 2003).

Most of the tools that have been developed for the assessment of e-commerce systems (Molla, 2001, Offut, 2002) give emphasis on the web applications of the system and they are based on surveys. This process provides significant results but demands extra time for data collection and data analysis in each measurement phase. The method's tool provides a flexible way to define the quality of e-commerce systems, as users perceive it, in a short period of time.

This paper presents a method for assessing the quality of e-commerce systems. This method is based on a previously presented model (Stefani, 2003) using Belief Networks. This model is used for the assessment of e-commerce systems (developed or ever during their development). The model offers numerical results for the overall quality of an e-commerce system, as well as for its intermediate factors and lower-level criteria. This paper presents the experimental results and the data analysis that aided towards the calibration of the model. Assessing an e-commerce system and its individual characteristics based on the numerical results were derived from the model.

In section 2 the foundations of the proposed method and the tool used are presented, while in section 3 the aim and the context of the study are discussed. In section 4 the experimental results that aided to the definition of the numerical scales are presented, while in section 5 the application of the presented method is further discussed. Finally, in section 6 conclusions and future work are presented.

2 PRESENTATION OF THE MODEL

The method presented in this paper uses a Belief Network in order to model the quality factors of e-commerce systems. This model is based on the ISO 9126 quality standard (ISO, 1991) and specifically it relies on the quality characteristics and

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 91–97.

sub-characteristics that are directly related to quality as perceived by the end-users. These quality characteristics are: Functionality, Usability, Reliability and Efficiency.

The mathematical model on which Bayesian Networks are based, is the theorem developed by the mathematician and theologian Thomas Bayes. The Bayesian Networks are a special category of graphic models where the nodes represent variables and the directed arrows represent the relation between the nodes. In the Bayesian Networks the node from which the directed arrow starts is defined as 'parent' node whereas the node where the directed arrow points at is defined as a 'child' node. Therefore, a Bayesian Network is a graphic network that describes the relations of probabilities between variables.

In order to define the relations between the variables, firstly the dependent probabilities that describe the relations between the variables must be determined for each node. If the values of each variable are distinct, then the probabilities for each node can be described in a Node Probability Table. This table presents the probability that a 'child' node is assigned a certain value for each combination of possible values of the 'parent' nodes. For example, if there is a Bayesian Network that presents one child node A and two parent nodes B, C then the probability table of node A reflects the probability P(A| B, C) for all possible combinations of A, B, C.

The Belief Network of the model consists of a number of nodes. Because of the hierarchical structure of the ISO 9126 quality standard, this Belief Network is represented as a tree. The root of this tree is the node Quality, which represents the e-commerce system quality as a whole. This node is connected to four nodes, one for each of the four aforementioned quality characteristics. Furthermore, each quality characteristic node is connected to the corresponding quality sub-characteristics, according to the ISO 9126. Finally, each of these quality sub-characteristic nodes is connected to intermediate nodes or to leaf nodes comprising the characteristics of e-commerce systems. A graphical presentation of a part of the network is illustrated in Figure 1.

The leaf nodes can be measured without subjectivity, since they simply answer the question posed to the user whether a specific e-commerce characteristic exists in the system or not. As a result, they take values 0 or 1. All the intermediate nodes are characterized by three possible states: 'good', 'average' and 'poor', except the central node (the Quality node), which is characterized by two possible states: 'good' and 'poor'. In this model all these possible states of each intermediate node take probability values which vary between 0 to 1. The probabilities of the model are based on data taken from previous studies of e-commerce systems (Stefani, 2001).

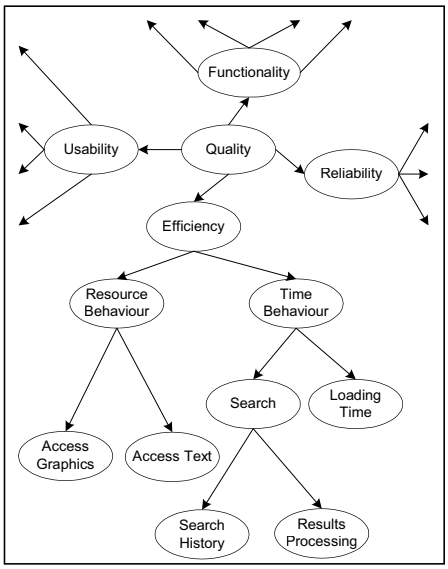


Figure 1: Graphical presentation of the model.

The use of the model can be forward and backward. In the forward use, the user inserts evidence to the leaf nodes in order to assess the overall quality of an e-commerce system. In this way, the model estimates the system's quality and characterizes it as 'good' and 'poor' also providing the corresponding probability values. Whereas the backward use of the model provides assessments regarding for intermediate or leaf nodes, when the value of a parent node is defined.

The model is also distinguished by its dynamic character. In other words, the node probability tables can always be refined by its use, while the results derived from its application can be utilized for its constant improvement, contributing to a continuous evolvement and upgrading. The current version of the model with simple instructions of its use is available on the web site of the Software Quality Research Group of the Hellenic Open University (SQRG-HOU, 2004).

3 AIM AND CONTEXT OF THE STUDY

The method proposed in this paper is based on the use of the aforementioned model. However, in order

for the application of the model to be worthwhile the probability values for its nodes must be meaningful. A comparative approach between these probability values and the assessment of each e-commerce system must be formulated. In other words, when the model estimates the probability values of the quality characteristics of an e-commerce system, one must be able to classify this system and ascertain the specific fields that need to be improved. The paper provides the boundaries and the scales of these values that were concluded from experimental measurements to a number of e-commerce systems. As a result, it provides a non-subjective way of characterizing an e-commerce system according to the quality characteristics with which this method is concerned. In this way, the use of this model can easily lead to conclusions and determine specific corrective actions needed to be set in order to improve the quality of the system.

In this case study a number of e-commerce systems were measured following the proposed method. In detail, the data of this method were based on the assessment of 120 different Business to Consumer (B2C) e-commerce systems. The selection of these sites was a representative list randomly sampled from the entire list of the Greek and international e-commerce systems, which were available the day of this study. For each of these systems the aforementioned model was used, by defining the values in its child nodes. In this way, the probability values for all the intermediate nodes were estimated.

After collecting the measurement data, the next step was the analysis of the results. Firstly, the normal distribution of the data for all the quality characteristics and sub-characteristics was checked in order to ensure their validity. As previously mentioned, the aim of this research was the determination of the possible boundaries and scales of the measurement data, so as to define in an easy and non-subjective way which quality characteristics of an e-commerce system have high or low scores. In other words, the measurement data must be grouped in different clusters that characterize how good or bad a system is. In order to define these clusters, three alternative approaches could be followed: a) setting a priori the values that define the boundaries of the clusters, even before conducting the experimental measurements to the number of e-commerce systems used in this research, b) setting these values with the use of percentages of the measurement results and c) estimating these values by judging from the measurement results themselves and their possible distribution to clusters.

Although all of the alternative approaches of analyzing the data are acceptable, the third one was chosen, since it provides more representative rates. Moreover, the analysis of the data showed that they were clearly distributed in different clusters. In this way the desired boundaries of the different scales for each quality characteristic were defined with more accuracy. Following this approach in the analysis, these boundaries can be defined regardless of the number of the e-commerce systems that were measured in this research.

The benefits of this analysis are noticeable. First of all, it provides an easy and non-subjective way to rank an e-commerce system according not only to the overall quality, but for each quality characteristic or sub-characteristic as well. In this way, it obviously shows which corrective actions may be followed to improve the quality of the system. Moreover, using this analysis, developers are able to determine the e-commerce characteristics on which they must focus, in order to achieve a desired value for the quality of the system that they develop.

4 DATA ANALYSIS

By applying the model at 120 e-commerce systems, measurement results were collected for the overall quality, the quality characteristics and sub-characteristics of the presented method. The measurement results, as presented hereinafter, correspond with the node probability values offered from the method's tool for the state "Good". The entire set of probability values for all states can be found in the web site of the Software Quality Research Group of the Hellenic Open University (SQRG-HOU, 2004).

The results for the overall quality were distributed normally and are presented in Table 1 *(Normality test Kolmogorov-Smirnov Significance level(Nom. test K-S s.l.)= 0.1; Mean (m)=0.56; Standard deviation(Std)= 0.24),* and they were distinguished in 3 categories A, B, C.

- Category A includes a small number of measurement values because in this category were included the e-commerce systems that satisfy strict criteria for the overall quality, the quality characteristics and sub-characteristics. And as the nature of the e-commerce system is to give emphasis to some of the above quality characteristics it is extremely difficult to achieve high measurement results in all sub-characteristics.
- Category B comprises e-commerce systems that satisfy a number of criteria for the overall quality. Although these systems appear to have an acceptable probability value for quality, corrective actions can be followed in order to

achieve a score as high as systems in category A have.

- In category C were placed the e-commerce systems that present low measurement results for the overall quality and the other components of the presented model.

Table 1: Measurement results for the overall quality.

0,1632	0,2006	0,2036	0,2073	0,2100	0,2116
0,2211	0,2216	0,2226	0,2299	0,2387	0,2426
0,2558	0,2560	0,2690	0,2745	0,2754	0,2813
0,2848	0,2905	0,2954	0,2995	0,3000	0,3044
0,3051	0,3164	0,3169	0,3240	0,3247	0,3284
0,3396	0,3405	0,3437	0,3519	0,3537	0,3561
0,3572	0,3575	0,3750	0,3849	0,3907	0,3910
0,3935	0,4000	0,4079	0,4372	0,4500	0,4526
0,4595	0,5033	0,5088	0,5530	0,5549	0,5606
0,5768	0,5802	0,5921	0,6000	0,6000	0,6008
0,6047	0,6106	0,6118	0,6153	0,6155	0,6155
0,6256	0,6258	0,6264	0,6369	0,6461	0,6500
0,6602	0,6696	0,6732	0,6747	0,6830	0,6849
0,6926	0,7064	0,7210	0,7310	0,7352	0,7370
0,7371	0,7410	0,7536	0,7550	0,7553	0,7553
0,7559	0,7643	0,7643	0,7660	0,7715	0,7759
0,7777	0,7790	0,7870	0,7884	0,7978	0,7978
0,8000	0,8203	0,8303	0,8404	0,8478	0,8599
0,8965	0,8974	0,9002	0,9059	0,9190	0,9201
0,9300	0,9542	0,9813	0,9813	0,9837	0,9860

The distribution of the measurement results in three categories was based on the clear separation, which their values present. Analytically the measurement results for the overall quality were distributed in three clusters as is presented in the histogram of Figure 2. Each cluster presented minimum and maximum values that were used to define the boundaries for the corresponding category. The boundaries of each category were defined by using the formula (1).

$$X = \left(X_{\max} + X_{\min}\right) / 2 \qquad (1)$$

Category A, for the overall quality of e-commerce systems comprises measurement values where $x > 0.89$, *(Nom. test Shapiro-Wilk s.l = 0.1; m = 0.93; Std = 0.136).* Category B comprises e-commerce systems that present $0.55 < x < 0.89$, *(Nom. test K-S s.l = 0.1; m = 0.69; Std = 0.185).* Finally category C comprises measurement values where $x < 0.55$ *(Nom. test K-S s.l = 0.2; m = 0.32; Std = 0.082),* as it is presented at the histogram in Figure 3.

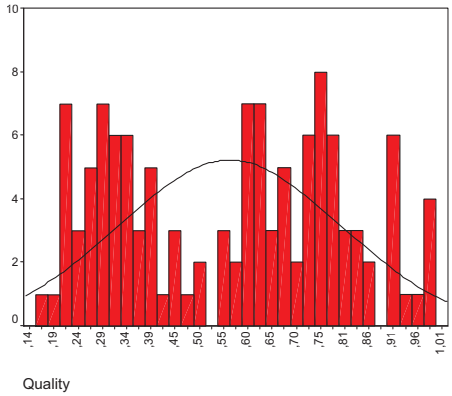


Figure 2: Histogram for the overall quality.

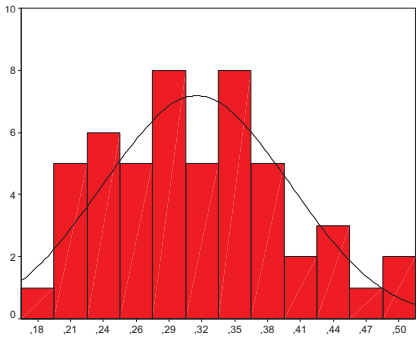


Figure 3: Histogram for category C of Quality.

This process has been applied for the quality characteristics and sub-characteristics of the model ¡n order to define the boundaries for each of them. It should be recorded that all measurement results for quality characteristics, and sub-characteristics, were distributed normally. For example the quality characteristic of Efficiency comprises measurement results *(Norm. test K-S s.l= 0.1; m= 0.59; Std.= 0.29)* that are presented in Figure 4 and were distributed also in three categories A,B,C.

Analytically, the scale calibration of the quality characteristics and sub-characteristics is presented in Table 2.

Table 2: Scale calibration of quality characteristics.

Scale Calibration			
Category	**A**	**B**	**C**
Quality	X > 0,88	0,88 > x > 0,53	X < 0,53
Functionality	X > 0,82	0,82 > x > 0,55	X < 0,55
Security	X > 0,82	0,82 > x > 0,55	X < 0,55
Interoperability	X > 0,93	0,93 > x > 0,80	X < 0,80
Suitability	X > 0,83	0,83 > x > 0,46	X < 0,46
Accuracy	X > 0,83	0,83 > x > 0,61	X < 0,61
Reliability	X > 0,84	0,84 > x > 0,62	X < 0,62
Fault Tolerance	X > 0,80	0,80 > x > 0,57	X < 0,57
Maturity	X > 0,80	0,80 > x > 0,62	X < 0,62
Recoverability	X > 0,84	0,84 > x > 0,62	X < 0,62
Usability	X > 0,87	0,87 > x > 0,63	X < 0,63
Attractiveness	X > 0,89	0,89 > x > 0,72	X < 0,72
Learnability	X > 0,90	0,90 > x > 0,60	X < 0,60
Understandability	X > 0,82	0,82 > x > 0,57	X < 0,57
Efficiency	X > 0,90	0,90 > x > 0,39	X < 0,39
Resource Behavior	X > 0,87	0,87 > x > 0,53	X < 0,53
Time Behavior	X > 0,86	0,86 > x > 0,44	X < 0,44

The scale calibration, which is represented in the table, comprises three categories A, B, C. Figure 5 presents a measurement for category B of Reliability (*Norm. test Shapiro Wilk s.ll = 0.454>>0.05, m= 0.73, Std. = 0,06)*.

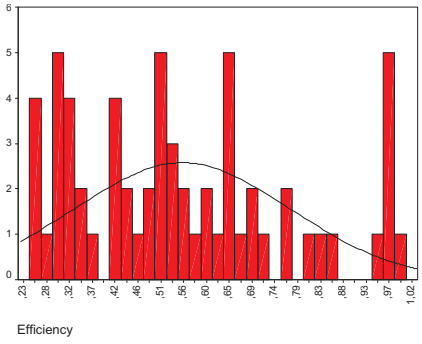


Figure 4: Histogram for the Efficiency.

In Figure 6 category A of Usability *(Norm. test Shapiro Wilk s.l = 0.443>>0.05, m = 0.93, Std. = 0.3)* is also presented.

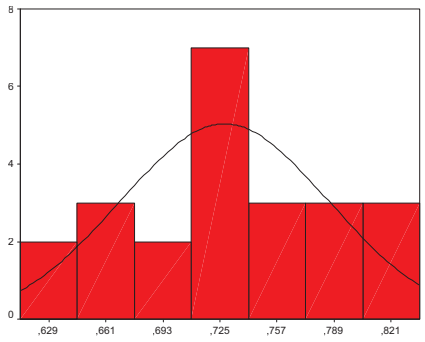


Figure 5: Histogram for category B of Reliability.

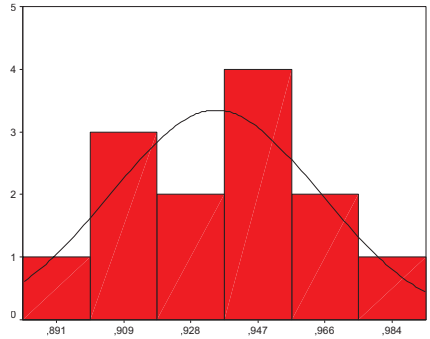


Figure 6: Histogram for category A of Usability.

The statistical analysis for the overall quality, the quality characteristics and sub-characteristics were offered in the quality's research group web site [SQRG-HOU, 2004].

5 THE PROPOSED METHOD

The proposed method's tool can be used for the assessment of e-commerce systems in order to identify problematic or high quality applications or modules. In detail, the tool provides probability values for the overall quality, the quality characteristics and sub-characteristics of an e-commerce system. The meaning of each value can be explained using the scale calibration table of the method. In other words, in the forward use of the method's model a user is able to give evidence to the leaf nodes in order to estimate the probability values of each quality characteristic. Afterwards, by means of Table 2, one can identify the cluster to which each quality characteristic belongs. So, it is

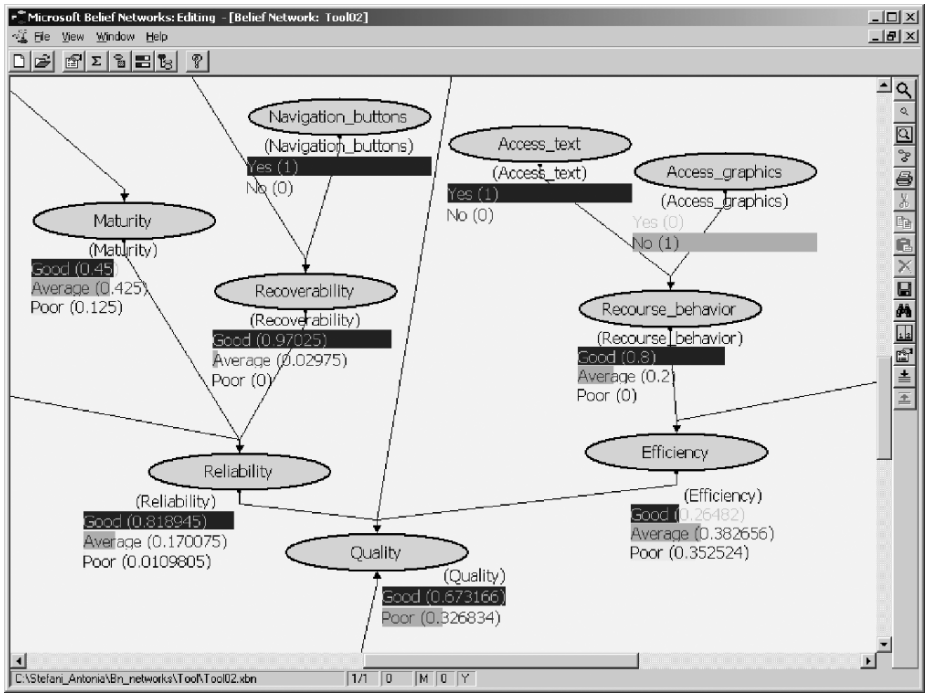


Figure 7: A screen shot of the method's tool.

possible to detect where an e-commerce system's drawbacks may exist in order to improve its quality. Additionally, it is possible to define what should be improved in an e-commerce system or where more emphasis should be given in the system.

The proposed method provides a non-subjective way of characterizing an e-commerce system according to the quality characteristics with which this method is concerned. The non subjective character of the model means that a simple user of this measurement tool is asked to define evidence in a binary way by providing a simple 'Yes' or 'No'. Only the developer of the model defines the values of the node probability table based on his/her experience and his/her specialized knowledge.

In detail, the use of the proposed method can be described as follows. For the leaf nodes the user gives as an evidence a simple yes or no, whether a specific e-commerce characteristic exists or not in the system. For example, in the case of leafs "FAQs" and "Shopping list", the user can easily define if they are available in the system or not. Knowing a priori the boundaries of each cluster of the quality, the mean value and the standard deviation, it is easy to define the overall quality of the system. Moreover, this process can be applied for each quality characteristic or sub-characteristic. In this way, one can identify which parts of the system need improvement. As a result, one is able to determine the specific corrective actions needed to be set in order to improve the quality of an e-commerce system.

This method can also be applied when developing a new e-commerce system. In this case, as the developers have already designed the characteristics and functions that the system will consist of, and know the preferences of the end-users of the system, they can focus on specific quality characteristics and improve them. Using the method's tool they must give more emphasis on the e-commerce characteristics that relay the quality characteristics they want. So, it could be possible to develop an e-commerce system that will be acceptable to end-users' quality requirements.

A screen shot of the method's tool is showed in Figure 7. This figure represents the probability values of an e-commerce system that has been used in the experiments of this research. Evidence has been inserted in the leaf nodes of the tool, so as to

measure the probability values of all the intermediate nodes.

6 CONCLUSIONS AND FUTURE WORK

This paper presents a method to assess the quality of e-commerce systems, which can also help developers of such systems during the design phase. It determines the boundaries and the scales of the probability values of the method's tool for all the quality characteristics. This determination, which is the main aim of the paper, was concluded from experimental measurements to a number of e-commerce systems and is presented in brief here.

The method's tool and all the experimental results derived from this research are available to whoever wishes to conduct similar measurements. Future work includes the collection of data from corresponding experiments performed by other researchers by means of this method. It also includes the application of the method during the design phase of an e-commerce system, and the analysis of the results, which will be derived from it.

REFERENCES

ISO, 1991. *Information technology - Evaluation of software - Quality characteristics and guides for their use.* International Standard: ISO/IEC 9126.

Lohse, G, Spiller, P, 1998. Electronic Shopping Designing online stores with effective customer interfaces, *Communications of the ACM.*

Molla, A, Licker, P, 2001. E-Commerce Systems Success: An Attempt to extend and respecify the Delone and Maclean model of IS success, *Journal of Electronic Commerce Research*

Nielsen, J, 2000. *Designing Web Usability, The Practice of Simplicity*, New Riders Publishing, Indianapolis.

Offutt, M, 2002. Quality Attributes of Web Software Applications, IEEE Software, March/April 2002

Schafer, B, Konstan, J, Riedl, J, 2001. E-Commerce Recommendations Applications, *Data Mining and Knowledge Discovery*.

Sherman, A, Karat, J, Karat, C, Brodie, C, and Vergo, J, 2003. User Attitudes Regarding a User-Adaptive e-Commerce Web Site, *User Modelling and User-Adapted Interaction.*

SQRG-HOU, 2004. Web site of the Software Quality Research Group of the Hellenic Open University, http://artemis.eap.gr/quality

Stefani, A, Xenos, M, 2001. A model for accessing the quality of e-commerce systems. *Panhellenic Conference with International Participationin Human Computer Interaction (PC-HCI-2001)*.

Stefani, A, Xenos, M, Stavrinoudis, D, 2003. Modelling E-Commerce Systems' Quality with Belief Networks. *In VECIMS 2003-International Symposium on Virtual Environments, Human-Computer Interfaces, and Measurement Systems.* IEEE

Wilson, L, 2003. Designing an electronic commerce interface: attention and product memory as elicited by web design. *Electronic Commerce Research and Application.*

GENDER DIFFERENCES IN ONLINE SHOPPERS' DECISION-MAKING STYLES

Chyan Yang and Chia Chun Wu
Institute of Business and Management, National Chiao Tung University, Sec. 1, Jhongsiao W. Rd, Taipei, Taiwan(ROC)
Email: professor_yang@hotmail.com, wuchiachun.bm90g@nctu.edu.tw

Keywords: Internet shopping, Decision-making styles, Gender differences, Exploratory factor analysis, Discriminant analysis.

Abstract: Because of the SARS epidemic in Asia, people chose to the Internet shopping instead of going shopping on streets. In other words, SARS actually gave the Internet an opportunity to revive from its earlier bubbles. The purpose of this research is to provide managers of shopping Websites regarding consumer purchasing decisions based on the CSI (Consumer Styles Inventory) which was proposed by Sproles (1985) and Sproles & Kendall (1986). According to the CSI, one can capture the decision-making styles of online shoppers. Furthermore, this research also discusses the gender differences among online shoppers. Exploratory factor analysis (EFA) was used to understand the decision-making styles and discriminant analysis was used to distinguish the differences between female and male shoppers. Managers of Internet shopping Websites can design a proper marketing mix with the findings that there are differences in purchasing decisions between genders.

1 INTRODUCTION

Taiwan's Internet users reached 8.76 million by June 2003, as reported by Institute for Information Industry ECRC-FIND Center. Compared with last year, the Internet users only grew by 90 thousands. This means that Taiwan's Internet market has become more mature gradually. In spite of the mature Internet market, there is seldom successful E-business and this phenomenon leads to the Internet bubbles.

Unfortunately the SARS epidemic broke out in spring 2003 in Asia. However, this crisis did give the slow Internet market a boost because people stayed at home whenever possible. In consideration of the chance to recover the prosperity, this research attempts to help marketing managers provide suitable marketing strategies. Therefore, this research used exploratory factor analysis to find consumers' decision-making styles by the CSI, which was proposed by Sproles (1985) and Sproles & Kendall (1986). By understanding the consumers' decision-making styles, managers of shopping Websites can hold more advantageous activities to arouse the consumers' interest and improve sales

In-store purchases account for the vast majority of consumer buying. Increased time pressure on either genders, especially on women, has been cited as one of the principal advantages of catalogue and online shopping. It has been broken gradually that the stereotype of an Internet shopper appears to be a youngish, well-educated man (Alreck & Settle, 2002). As reported by Nielsen//NetRatings, there are 35 millions of female internet users in Europe, which is almost 42% of European Internet users. Moreover, concerning the ranking of the main countries in the World, the percentage of American female Internet users is 51%, and the highest and it's about 51%. In Sweden and UK, the proportions of female Internet users are both over 45%. Other counties such as Netherlands, France, Switzerland, Spain and German are all over 40%. The report also shows that shopping, travelling, education, finance, health, and beauty care Websites are the most attractive to female Internet users (Institute for Information Industry, ECRC-FIND).

The same phenomenon can also be found in Asia-Pacific region. Female Australian Internet users are 48% of the whole Australian Internet users, 46% of New Zealand, 45% of South Korea, 44% of Hong Kong, 42% of Singapore, and 41% of Taiwan. Among these countries, the growth of South Korea female Internet users is the fastest, which rate reaches 55%. The rest are Taiwan (27%), Singapore (16%), Australia (16%), and Hong Kong (11%). New Zealand is 10%, which is the lowest growth

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 99–106.

rate (Institute for Information Industry, ECRC-FIND).

2 LITERATURE REVIEW

2.1 Decision-Making Style

A consumer decision-making style is defined as a mental orientation characterizing a consumer's approach to making choices. It has cognitive and affective characteristics (Sproles & Kendall, 1986). Extant research in this field has identified three approaches to characterize consumer styles: (1) the Consumer Typology Approach; (2) the Psychographics/Lifestyles Approach; and (3) the Consumer Characteristics Approach. The Consumer Characteristics Approach is one of the most promising as it deals with the mental orientation of consumers in making decisions (Durvasula, Lysonski, and Andrews, 1993).

The original of this approach was based on an exploratory study by Sproles (1985) that identified fifty items related to this mental orientation. Afterward, Sproles & Kendall (1986) reworked this inventory and developed a more parsimonious scale with forty items (Durvasula, Lysonski, and Andrews, 1993). These items were titled Consumer Style Inventory. Many studies that discussed consumer decision-making style refered to Sproles (1985) and Sproles & Kendall (1986) as the base. Some relative studies were shown as Table 1.

2.2 Gender Differences in Internet

There have been many studies which contribute to gender differences in the application of Internet. Gefen & Straub (1997) extended the Technology Acceptance Model to IT diffusion and used this structure to discuss gender differences in the perception and use of E-Mail. They found that gender differences indeed influenced the use of E-Mail. Jackson, Ervin, Gardner & Schmitt(2001) used path analysis to discuss the use of Internet between the two genders and found some influential factors such as motivational, affective and cognitive factors. The results were shown that women used Internet as a communication tool while men used it as a search tool.

Boneva, Kraut & Frohlich(2001) discovered that women used E-Mail as a personal relationship tool more than men did. Furthermore, Teo & Lim(1997) investigated 1370 Singapore residents. They used Internet to understand the gender gap about usage patterns and perception of the Internet. The result has important implication for business who seeks to sell products targeted at female consumers via the Internet. The reason is female are well-educated.

Based the above studies, we added gender difference in consumer decision-making styles. There must be some differences while online shoppers make decisions because Internet shopping is a kind of application of Internet.

3 METHODOLOGY

3.1 Questionnaire Design

Translation was used to prepare the forty-item CSI scale for the investigation because of the language and culture in Taiwan. Slight changes must do owing to the purpose of this research, for example, we added such words like "online shopping" in the items. A five-point scale was used, ranging from strongly disagree to strongly agree. Moreover, we used Internet questionnaire instead of traditional one. The reason was lain on convenience and time-saving to use this kind of method to delivery questionnaire.

3.2 Sample Selection

Convenient sampling of 209 Internet users that consisted of 102 females and 107 males is conducted. Besides, all these 209 responses were from those with Internet shopping experiences. For the sake of deciding online shoppers' decision-making styles, this research used exploratory factor analysis (EFA). Although there were many researches that discussed CSI, none used CSI to online shopper. Additionally, we contested that the gender differences might lead to different decision-making styles. The method we adopt to recognize genders differences is discriminant analysis. EFA and discriminant analysis were tested by using SAS 8.2, and results were shown next section.

4 RESULTS

4.1 Reliability and Validity

In social science research, one of the most widely-used indices of internal consistent reliability is Cronbach Alpha (Cronbach, 1951). It can save time to measure the reliability comparing with test-retest reliability and it's measurement effect is as well as test-retest reliability. A widely-used rule of the

thumb of 0.7 has been suggested by Nunnally (1978). Reliability coefficient in this research is more than 0.7(Cronbach coefficient alpha = 0.86), so the questionnaire we used has internal consistent reliability. Besides internal consistent reliability, we should consider the validity of the questionnaire. The questionnaire possessed content validity because we adopted from CSI which was suggested by Sproles (1985) and Sproles & Kendall (1986).

4.2 Results of Exploratory Factor Analysis

An exploratory factor analysis (EFA) was performed to categorize online shoppers' decision-making styles. Consistent with Sproles & Kendall (1986), principal components analysis with varimax rotation was used. Because principal components analysis didn't produce a single solution but left the decision about the right number of factors largely to researchers, we chose eigenvalue-one as criterion to decide the number of factors (Kaiser, 1960). The rule of eigenvalue-one is that the number of factors is decided when eigenvalue is greater than one. This research we classified seven factors (Table 2). The results of EFA were shown in Table 3.

Factor 1: Perfectionism
This kind of online shopper values the quality of products. When it comes to purchasing products, they try to get the very best or perfect choice. In general, they usually try to buy the best overall quality.

Factor 2: Novel-Fashion Consciousness
This kind of online shopper likes to buy the fashionable and novel goods. They are the early adopter. They keep their wardrobe up-to-date with the changing fashions. Fashionable, attractive styling is very important to them.

Factor 3: Price Consciousness
This kind of online shopper very considers the value of money. The lower price products are usually their choice. They usually take the time to shop carefully for best buys

Factor 4: Confused by Overchoice
This kind of online shopper is worry about much information about products. Too much information will disturb them to make right purchase decisions. The more they learn about products, the harder it seems to choose to best. All the information they get on different products confuses them.

Factor 5: Brand Consciousness
This kind of online shopper values the brand of products. The well-known national brands are best for them to choose. They think the more expensive brands are usually their choice.

Factor 6: Recreational Shopping
This kind of online shopper thinks shopping will waste time unless it can please him. A product doesn't have to be perfect, or the best, to satisfy them. They enjoy shopping just for the fun of it.

Factor 7: Brand-Loyal Consciousness
This kind of online shopper is brand loyalist. They have favorite brands they will buy over and over. Once they find a product or brands they like, they will stick with it.

4.3 Results of Discriminant Analysis

First, we should test if the means have significant differences between seven factors in two populations (female and male) by one-way MANOVA before discriminant analysis. The result shows that seven factors' mean have significant differences between two populations (Wilks' Lambda = 0.86, F = 4.52, p = 0.0001, see Table 4).

Second, we chose the factors by stepwise discriminant analysis that could obviously discriminant difference between female and male. The result suggested that only Factor1, Factor 2, Factor 3 and Factor 5 could differentiate female from male.

Finally, we used Factor1, Factor 2, Factor 3 and Factor 5 to implement discriminant analysis. This research only had two populations, so there was only one discriminate function

$L = -0.3104F1 - 0.9435F2 + 0.5004F3 + 0.8142F5$. The standardized canonical coefficients are shown in Table 5. The total classification error rate is 0.4070, and the classification results are list in Table 6. This error rate means that we can classify correctly by this discriminant function and its correct rate is about sixty percentages. From the discriminate function, we can obtain discriminate scores. If the scores are higher than total mean, then it would be males' decision-making. If the scores are lower than total mean, then it would be females' decision-making. In general, it exists differences between female and

Table 1: Relative Research on Consumers' Decision-Making Styles.

Researchers	Sample Structure	Decision-Making Styles
Sproles 1985	A sample of 111 undergraduate women in two classes of the School of Family and Consumer Resources, University of Arizons	Six Decision-Making Styles: 1. Perfectionism 2. Value conscious 3. Brand consciousness 4. Novelty-fad-Fashion consciousness 5. Shopping Avoider 6. Confused, support-seeker style
Sproles Kendal 1986	482 students in 29 home economics classes in five high schools in the Tucson area	Eight Decision-Making Styles: 1. Perfectionistic, high-qualityconscious 2. Brand conscious 3. Novel-fashion conscious 4. Recreational, hedonistic consumer 5. Price conscious 6. Impulsive, careless consumer 7. Confused by overchoice consumer 8. Habitual, brand-loyal consumer
Hafstrom, Chae Chung 1992	310 college students at four universities in Taegu	Eight Decision-Making Styles: 1. Brand conscious 2. Perfectionistic, high-quality conscious 3. Recreational-shopping consumer 4. Confused by overchoice consumer 5. Time-engery conserving consumer 6. Impulsive, careless consumer 7. Habitual, brand-loyal consumer 8. Price-value conscious
Durvasula, Lysonsk Andrews 1993	210 undergraduate business students at a large university in New Zealand	Eight Decision-Making Styles: 1. Perfectionistic, high-quality conscious 2. Brand conscious 3. Novel-fashion conscious 4. Recreational, hedonistic consumer 5. Price conscious 6. Impulsive, careless consumer 7. Confused by overchoice consumer 8. Habitual, brand-loyal consumer
Jessie X. Fan Jing J. Xiao 1998	271 undergraduate students from Zhongshan University, South China Normal University, South China University of Technology, Guangdong Commercial College and Jinan University	Five Decision-Making Styles: 1. Brand consciousness 2. Time consciousness 3. Quality consciousness 4. Price conscious 5. Information utilization
Gianfranco Walsh, Vincent-Wayne Mitchell & Thorsten Hennig-Thurau(2001)	455 male and female shoppers who are entering or leaving a shop in Lünegurg and Hamburg	Seven Decision-Making Styles: 1. Brand consciousness 2. Perfectionism 3. Recreational/hedonistic 4. Confused by overchoice 5. Impulsiveness Price conscious 6. Novel-fashion consciousness 7. Varity seeking

Researchers	Sample Structure	Decision-Making Styles
Alice S. Y. Hiu, Noel Y. M. Siu, Chaile C. L. Wang & Ludwig M. K. Chang(2001)	387 consumer who are in shopping malls or places nearby shopping center in Guangzhou, China	Seven Decision-Making Styles: 1. Perfectionistic, high-quality 2. Brand conscious 3. Novel-fashion conscious 4. Recreational/hedonistic 5. Price conscious 6. Confused by overchoice 7. Habitual, brand-loyal consumer
Cathy Backwell & Vincent-Wayne Mitchell(2003)	244 female undergraduate students aged between 18 and 22	Five Decision-Making Styles: 1. Recreational quality seeker 2. Recreational discount seeker 3. Shopping and fashion uninterested 4. Trend setting loyal 5. Confused time/money conserve

male's decision-making style. Figure 1 shows the differences between two populations.

5 CONCLUSIONS

According to the CSI, online shoppers could be categorized into seven main decision-making styles: perfectionism, novel-fashion consciousness, price consciousness, confused by overchoice, brand consciousness, recreational shopping and brand-loyal consciousness. Compared with the findings of Sproles & Kendal 1986, online shoppers lack of the type of "impulsive careless consumer". This means that online shoppers are programmed problem solving while making purchase decisions. When people adapt online shopping, it means that they have already thought it carefully and might get used to shopping through Internet. Therefore, consumers in cyberspace and reality environment may act differently to some degrees.

Secondly, this research also discussed the gender differences among online shoppers. Discriminant analysis was employed to distinguish the differences between female and male shoppers. We discovered that female and male indeed exhibited some difference on decision-making styles from the discriminate function. Males are dominated over price consciousness and brand consciousness and females are dominated over perfectionism and novel-fashion consciousness. Meanwhile, these findings can provide managers of Internet shopping Websites to design a proper homepage and marketing mix for males and females.

Third, further researchers can use the seven online shoppers' decision-making styles as segmentation variables to capture more details about online shoppers. This research can propose some aspects for both researchers and practitioners who are interested in consumer behavior in E-Commerce.

REFERENCES

Taiwan Regular Internet Users Reached 8.76 Millions By June 2003, 2003/8/15, http://www.find.org.tw /0105/ howmany/howmany_disp.asp?id=57, Institute for Information Industry ECRC-FIND Center (in Chinese)

American Female Internet Users Are More Than Male and Female Internet Users Has Grown Rapidly in Asia-Pacific Region, 2001/8/31, http://www.find.org.tw /0105/property/0105_property_disp.asp?board_id=24, Institute for Information Industry ECRC-FIND Center (in Chinese)

European Female Internet Users Reaches 35 Millions, 2003/7/1, http://www.find.org.tw/0105/news/0105_ news_disp.asp? news_id=2737, Institute for Information Industry ECRC-FIND Center (in Chinese)

Alreck, Pamela & Settle, Robert B., 2002, Gender Effects on Internet, Catalogue and Store Shopping, *Journal of Database Marketing & Customer Strategy Management*, Jan, 9, 2, 150-162.

Bakewell, Cathy & Mitchell, Vincent-Wayne, 2003, Generation Y Female Consumer Decision-Making Styles, *International Journal of Retail & Distribution Management*, 31(2), 95-106.

Boneva, Bonka, Kraut, Robert & Frohlich, David, 2001, Using E-Mail for Personal Relationships: the

Table 2: The Criterion to Decide Factor Numbers.

	Eigenvalue	Difference	Proportion	Cumulative
1	6.50897283	3.20199722	0.3178	0.3178
2	3.30697562	0.91748347	0.1615	0.4793
3	2.38949214	0.45218978	0.1167	0.5960
4	1.93730236	0.31031516	0.0946	0.6906
5	1.62698720	0.40328094	0.0794	0.7701
6	1.22370625	0.18305716	0.0598	0.8298
7	1.04064909	0.10621954	0.0508	0.8806

Table 3: Taiwan Online Shoppers' Style Characteristics: Seven-Factor Model (wordings are directly adopted from Sproles (1985) and Sproles & Kendall (1986)).

Factor	Items	Factor Loadings
Factor 1	1 Getting very good quality is very important to me.	0.74
	2 When it comes to purchasing products, I try to get the very best or perfect choice.	0.83
	3 In general, I usually try to buy the best overall quality.	0.86
	4 I make special effort to choose the very best quality products.	0.74
	6 My standards and expectations for products I buy are very high.	0.60
Factor 2	15 I usually have one or more outfits of the very newest styles.	0.51
	16 I keep my wardrobe up-to-date with the changing fashions.	0.75
	17 Fashionable, attractive styling is very important to me.	0.79
	18 To get variety, I shop different stores and choose different brands.	0.69
	19 It's fun to buy something new and exciting.	0.52
Factor 3	24 I make my shopping trips fast.	0.54
	25 I buy as mush as possible at sale prices.	0.54
	26 The lower price products are usually my choice.	0.60
	31 I take the time to shop carefully for best buys.	0.61
	32 I carefully watch how mush I spend.	0.55
Factor 4	34 Sometimes it's hard to choose which stores to shop.	0.48
	35 The more I learn about products, the harder it seems to choose to best.	0.83
	36 All the information I get on different products confuses me.	0.82
Factor 5	9 The well-known national brands are best for me.	0.68
	10 The more expensive brands are usually my choice	0.75
	11 The higher the price of a product, the better its quality.	0.54
Factor 6	5 I usually don't give my purchases much thought or care.	0.48
	7 I shop quickly, buying the first product or brand I find that seems good enough.	0.41
	8 A product doesn't have to be perfect, or the best, to satisfy me.	0.50
	23 I enjoy shopping just for the fun of it.	0.49
Factor 7	37 I have favorite brands I buy over and over.	0.76
	38 Once I find a product or brands I like, I stick with it.	0.77

Table 4: Multivariate Analysis Results.

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.86407272	4.52	7	201	0.0001
Pillai's Trace	0.13592728	4.52	7	201	0.0001
Hotelling-Lawley Trace	0.15731001	4.52	7	201	0.0001
Roy's Greatest Root	0.15731001	4.52	7	201	0.0001

Table 5: Standardized Canonical Coefficients.

Variable	Can1
F1	–.3104382119
F2	–.9434892732
F3	0.5004005177
F5	0.8141717820

Table 6: Classification Results.

Predicted Group / Actual Group	Female	Male	Total
Female	59 (57.84%)	43 (42.16%)	102 (100%)
Male	42 (39.25%)	65 (60.75%)	107 (100%)
Total	101 (48.33%)	108 (51.67%)	209 (100%)

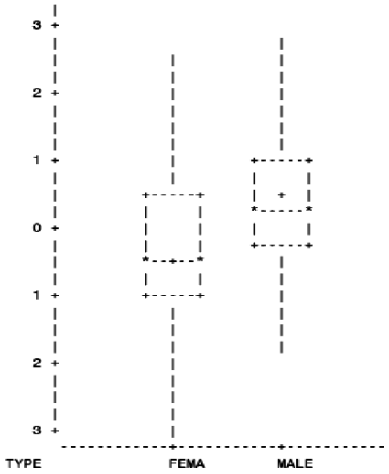


Figure 1: Gender Differences in Decision-Making Styles (Show by box-and-whisker plot).

Difference Gender Makes, *The American Behavioral Scientist*, Nov, 45, 3, 530-549.

Briones, Maricris G., 1998, On-line Retailers Seek Ways to Close Shopping Gender Gap, *Marketing News*, Sep 14, 32, 19.

Durvasula, Srinivas, Lysonski, Steven & Andrews, J. Craig , 1993, Cross-Culture Generalizability of a Scale for Profiling Consumers' Decision-Making Styles, *Journal of Consumer Affairs*, 27(1), 55-65.

Fan, J. X. & Xiao, J. J., 1998, Consumer Decision-Making Styles of Young-Adult Chinese, *Journal of Consumer Affairs*, 32, 275-294.

Gefen, David & Straub, Detmar W., 1997, Gender Difference in the Perception and Use of E-Mail: An Extension to the Technology Acceptance Model, *MIS Quarterly*, Dec. 21, 4, 389-400.

Hafstrom, Jeanne L. , Chae, J. S., 1992, Consumer Decision-Making Styles: Comparison between United States and Korean Young Consumers, *Journal of Consumer Affairs*, 26(1), 146-158.

Hiu, Alice S.Y., Siu, Noel Y. M., Wang, Charlie C.L., & Chang Ludwig M. K., 2001, An Investigation of Decision-Making Styles of Consumers in China, *Journal of Consumer Affairs*, 35(2).

Jackson, Linda A., Ervin, Kelly S., Gardner, Philip D. & Schmitt, Neal, 2001, Gender and the Internet: Women Communication and Men Searching, *Sex Role*, Mar, 44, 5/6, 363-379.

Kaiser, H. F., 1960, The Application of Electronic Computers to Factor Analysis, *Educational and Psychological Measurement*, 20, 141-151.

Nunnally, J., 1978, *Psychometric Theory*, New York: MaGraw-Hill.

Sproles, G. B.□Kendall, E. L., 1986, A Methodology for Profiling Consumers' Decision-Making, *Journal of Consumer Affairs*, 20(2), 367-379.

Sproles, G. B., 1985, From Perfectionism to Fadism: Measuring Consumers' Decision-Making Styles, in *Proceedings of American Council on Consumer Interest*, pp. 79-85.

Teo, Thompson S. H. & Lim, Vivien K.G., 1997, Usage Patterns and Perceptions of the Internet: the Gender Gap, *Equal Opportunities International*, 16, 6/7, 1-8

Walsh, G., Mitchell, Vincent-Wayne, & Hennig-Thurasu, Thorsten, 2001, German Consumer Decision-Making Styles, *Journal of Consumer Affairs*, 35, 73-99.

DESIGN AND EVALUATION OF THE HOME NETWORK SYSTEMS USING THE SERVICE ORIENTED ARCHITECTURE

Hiroshi Igaki, Masahide Nakamura and Ken-ichi Matsumoto
Graduate School of Imformation Science, Nara Institute of Sciense and Technology, 8916-5 Takayama-cho, Ikoma, Nara Japan
Email: hiro-iga@is.naist.jp, masa-n@is.naist.jp, matumoto@is.naist.jp

Keywords: Web Services, Service-oriented architecture, Home network, distributed system.

Abstract: In the conventional home network systems (HNS), a powerful centralized server controls all electric home appliances connected to provide value-added integrated services. However, when the number of the appliances increases and the appliances become more sophisticated, the conventional architecture would suffer from problems in superfluous resources, flexibility, scalability and reliability. This paper proposes alternative architecture for HNS, which exploits the service-oriented architecture with Web Services. In the proposed architecture, each appliance is controlled by a Web service in a de-centralized manner. Then, the services autonomously collaborate with each other to achieve the integrated service scenarios. To evaluate the HNS at the design process, we also present four kinds of evaluation metrics: reliability, load, complexity, and coupling. Using these metrics, we conduct a comparative study among the proposed and the previous HNS architectures.

1 INTRODUCTION

Recent advancement in computer network technology enables electric home appliances to be connected in a network. The appliances, such as an air-conditioner, door sensors, lights, a TV and a DVD player, are connected with each other. The system consisting of such networked home appliances is generally called a *Home Network System* (HNS for short). Several commercial HNS products are already on the market (e.g., LG E, 2004; Samsung, 2004; Hitachi, 2004).

The appliances in HNS are controlled together to provide *integrated services*, which add more value and convenience to the daily life of home users. Typical integrated services include;

If the user comes home, the lights and the air-conditioner are automatically turned on.

When the user starts to watch DVD movies, the lights becomes dark and the volume on the TV is adjusted.

The current HNS mainly adopts the *server centralized architecture* (we call it SCA in the following), where a powerful and intelligent server (called *Home Server*) controls all the dumb appliances connected. In general, each appliance does not have advanced intelligence, and it just receives (a sequence of) commands from the server with a low-level and light-weight network adapter.

Since SCA is quite simple architecture, it is relatively easy to apply SCA to the HNS consisting of the *conventional* home electric appliances. However, in the near future, the SCA-based HNS will be faced with the following problems.

Since SCA generally requires proprietary middle-ware, it is difficult to achieve the interoperability among products from different vendors.

All the appliances heavily rely on the centralized Home Server. Therefore, the server suffers from the scalability problem when the number of appliances becomes large. Also, the server requires considerably high reliability, because all the integrated services stop when the server fails.

Even if the appliances come to have more intelligent processors and network devices, the HNS cannot make flexible use of the resources as the Home Server takes the main control. Thus, the quality of the integrated services is limited to the features implemented in the server.

This paper presents alternative architecture for HNS. Specifically, we propose to apply the *service oriented architecture* (SOA, for short) (Hao, 2003)

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 107–114.

with Web services to HNS. SOA is basically architecture to integrate distributed self-contained services using loose coupling and well-defined interfaces. In this paper, we assume the next-generation home electric appliances, which are intelligent enough to process Web service transactions with own processors and network devices.

Our key idea is to export features of each appliance as methods of Web service, and to make the features directly available from other appliances in an open and standard manner (i.e., SOAP/XML). Thus, the appliances can autonomously collaborate with each other to build the integrated services in the HNS. Since the proposed SOA-based HNS does not require any centralized server, it is expected to be more scalable and fault-tolerant. Also, more sophisticated and flexible integrated services can be developed.

In this paper, we conduct the architectural design of a practical HNS example using the SOA framework. Then, we propose a graph-based method to evaluate the design quantitatively, from the viewpoints of reliability, workload, functional complexity and coupling. These methods are applied to two different HNSs with SOA and SCA, in order to see the difference.

2 SERVICE-ORIENTED ARCHITECTURE (SOA) AND WEB SERVICES

SOA is an architectural style whose goal is to achieve loose coupling among interacting autonomous software agents. A *service* is a unit of tasks done by a service provider to achieve desired end results for a service consumer. The *interface* of a service is strictly typed so as to be processable by software agents of the service consumers. Through the interface, features of the service are exported to the network as methods. Since the interface is supposed to be unchanged, the consumer can use the service from a remote place, as if it were just an ordinary method invocation, without knowing internal logic or protocol message formats. This is known as the *loose coupling*. Using this concept, a service can autonomously collaborate with other services, which enables more sophisticated integrated services.

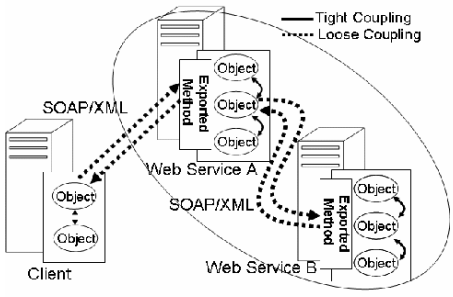


Figure 1: Service-Oriented Architecture.

A Web service (Ethan, 2002; W3C, 2004) provides an open and standard means to implement the SOA-based system. The interface of a Web service is described by XML-based format, specifically WSDL. Other systems interact with the Web service in a manner prescribed by its description using SOAP-messages, typically conveyed using HTTP with an XML serialization in conjunction with other Web-related standards.

Figure 1 shows an example of SOA using Web service. The client application (Client) accesses the first Web Service A through its exported method. Web Service A internally calls a method of Web Service B. Web Service B returns the result to Web Service A, and finally Client gets the end result. The interface of the exported methods is described by WSDL, and Client and Web Services are loosely coupled by SOAP/XML. As a result, Client uses the integrated service consisting of Web Service A and Web Service B (depicted by a large oval in the figure).

3 DESIGNING HOME NETWORK SYSTEM (HNS) WITH SOA

3.1 Key Idea

Considering today's evolution of network technology, it is reasonable to assume that the next-generation home electric appliances can be autonomous nodes with software control, supported by own processors and network devices (DHWG, 2004).

Our key idea is to apply SOA to such autonomous home electric appliances. Specifically, each appliance has a software layer (we call it *service layer*) from which its end device (hardware) can be controlled. Then, we implement an interface of the control in the service layer as a Web service, and export it to the network. By doing this, multiple appliances can autonomously collaborate with each other at the service layer. This enables to develop more interoperable and flexible integrated HNS services.

For instance, suppose that a Web service of room lights provides "SwitchON" method to the network. Then, a door sensor can collaborate with the lights by executing the method, so that the lights are turned on when the user opens the door. Note that this integrated service does not require any centralized server. Also, the communication between the sensor and the lights is done in terms of a standardized manner of Web services.

In the following subsections, we demonstrate how a practical HNS can be designed based on the proposed architecture with SOA and Web services.

3.2 Target Home Network System

As a practical example, in this paper, we try to design an HNS consisting of the following 9 home electric appliances: a DVD player, a TV, a speaker, a light, an illuminometer, a door, a telephone, an air-conditioner and a thermometer. In this HNS, we achieve the following eight service scenarios (denoted by SS) as the integrated services. These scenarios are taken from actual commercial products (ECHONET, 2004; Samsung, 2004).

SS1: The brightness of the light is automatically adjusted based on the current intensity of illumination with the illuminometer.
SS2: If the user enters a room from the door, the light are turned on.
SS3: When the user turns on the DVD player, the light becomes dark. Then, the TV and the speaker start in the DVD mode.
SS4: When the user watches the TV, the speaker is turned on.

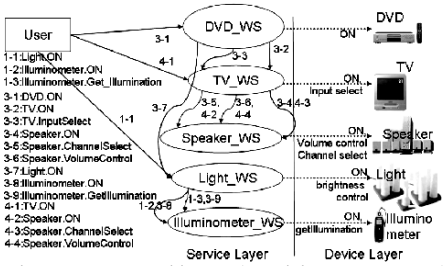


Figure 2: An HNS with SOA (containing SS1, SS3, SS4).

SS5: While the user is watching the TV, if the telephone rings, then the volume of the speaker becomes small.
SS6: The air-conditioning is optimized based on the thermometer.
SS7: If the user enters the room, the air-conditioner starts and adjusts the temperature to a comfortable degree.
SS8: When the user goes out or goes to bed, all the appliances are shut down and the door is securely locked up.

3.3 Design of the HNS with SOA

As discussed in Section 3.1, we assume that each appliance is an autonomous intelligent node, which can control the end device by the software. Also, each appliance is supposed to have enough processing power to operate own Web service to export its control to the network.

With the assumption, we here try to conduct an architectural design of the target HNS in Section 3.2 with SOA. Specifically, we consider what methods must be implemented in the Web service (denoted by WS, in the following) of each appliance, in order to achieve all the service scenarios SS1 to SS8 in the target HNS.

Figure 2 shows an architectural design involving a part of the service scenarios (SS1, SS3 and SS4). Each appliance consists of an end device and the corresponding WS (depicted by an oval). The whole architecture is divided into two layers: the *device layer* and the *service layer*. In the device layer, an end device is controlled directly by the corresponding WS (drawn by a dotted line). On the other hand, in the service layer, features of each appliance are exported as methods of the corresponding WS.

To provide the integrated service scenarios, the appliances collaborate with each other via the network, by autonomously executing the exported methods. In Figure 2, a solid arrow with label L from WS A to B means that WS A executes (uses) the method L provided by B. Due to the limited space, each label is represented by a number in the form of i-j describing j-th method executed in SS_i (I = 1,3,4).

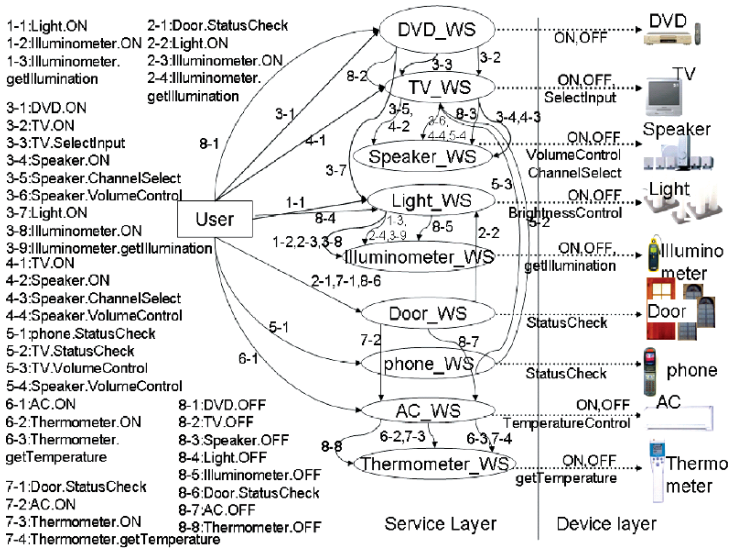


Figure 3: An HNS with SOA.

Let us consider SS1 in Section 3.2. In Figure 2, we can see a possible design to implement SS1, by traversing arrows prefixed by "1-". First, the user calls the method Light.ON to Light_WS. Next, Light_WS turns on the illuminometer with Illuminometer. ON, and acquires the current illumination by Illuminometer. Get Illumination. Finally, Light_WS sets up the optimal lighting to the lighting devices based on the present illumination. Similarly, by traversing the arrows prefixed by "4-", we can see the scenario SS4, where the TV autonomously turns on the speaker and adjusts the speaker volume.

SS3 can be achieved by reusing SS1 and SS4. The user first turns on the DVD player by DVD.ON. Next, DVD_WS executes TV.ON and TV.InputSelect for TV_WS, and calls Light.ON for Light_WS. Then, Light_WS and TV_WS execute the existing SS1 and SS4, respectively. Thus, the SOA allows us to reuse and integrate the existing scenarios to achieve a new integrated service.

An architectural HNS design containing all the scenarios SS1 to SS8 is shown in Figure 3

The main characteristics of the SOA-based HNS are summarized as follows. The distributed appliances collaborate with each other to realize the integrated service scenarios on demand from the user. Each WS can be used as a reusable component to construct integrated service scenarios. Since the control of HNS is fully distributed, the implementation of each WS is expected to be simple and self-contained.

3.4 Design of the HNS with SCA

For the comparison purpose, we also consider the target HNS with the server centralized architecture (SCA), which is adopted by most of the current commercial HNS products (Hitachi, 2004; Samsung, 2004). In these products, a light-weight adapter is connected to the conventional home electric appliance. The adapter relays commands from the powerful centralized server, called Home Server (HS, for short). The communication is performed by the proprietary software and protocol.

Figure 4 shows the architectural design of HNS with SCA. In this HNS, all the appliances are directly controlled by the HS. The integrated service scenarios are performed by cooperation of (tightly-coupled) objects in the HS. For example, when the user demands to execute SS3, Home Server directly sends the proprietary commands to the DVD player, the TV, the speaker and the lights.

Thus, in the SCA-based HNS, the architecture itself is quite simple since HS takes the control of all appliances. However, the implementation of HS tends to be complex, and the workload of the HNS is concentrated in HS. Also, all the integrated services become unavailable if HS fails. These issues are discussed quantitatively in the next section.

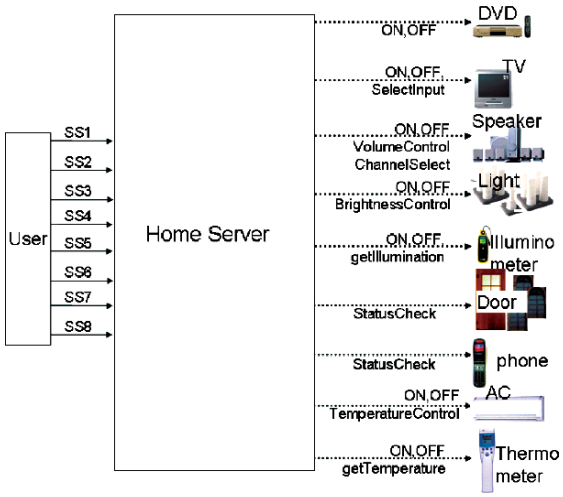


Figure 4: An HNS with SCA.

4 EVALUATION OF HNS ARCHITECTURAL DESIGN

In this section, we propose a graph-based method to perform quantitative evaluation of the HNS architectural design. For a given HNS design, the proposed method derives four kinds of metrics: reliability, workload, functional complexity and coupling.

4.1 Service Integration Graph

As seen in Figure 2, Fig. 3 and Figure 4, an HNS with integrated service scenarios (we simply call *scenarios* in the following) can be characterized by a labelled directed graph, where a node represents an HNS component (i.e., a user, an end device, a WS or an HS), and a directed edge denotes a method invocation among the components. By utilizing the graph, several important characteristics of the HNS can be mathematically derived.

A *labelled directed graph* G is defined by $G=(N,L,E)$, where N is a set of nodes, L is a set of labels, and $E \subseteq N \times L \times N$ is a set of labelled directed edges. For a given scenario s, a labelled directed graph $G=(N,L,E)$ is called a *service integration graph* for s, denoted by $SIG(s)$, iff G satisfies the following conditions:

- N is a set of all components appearing in s
- L is a set of all methods appearing in s
- An edge (p, m, q) exists in E iff p uses method m that is provided by q.

Next, we extend the service integration graph to the *set* of scenarios. Let s_1, s_2,..., s_k be a given set of scenarios. For each i ($1 \le i \le k$), we have $SIG(S_i)=(Ns_i, L\ s_i, Es_i)$. Then, we define $SIG(s_1, s_2,\ldots, s_k) = (\cup_i Ns_i, \cup_i Ls_i, \cup_i Es_i)$. If s_1, s_2,..., s_n are all the scenarios in the HNS, then we call $SIG(\{s_1, s_2,\ldots, s_n\})$ a *full service integration graph*, which is denoted by *FSIG*. Note that for a given HNS, any *SIG* is a subgraph of *FSIG*.

For instance, consider the scenarios SS1 to SS8 in Section 3.2. We can see that Figure 2 represents $SIG(\{SS1,SS3,SS4\})$ and that Fig. 3 represents $FSIG$ ($=SIG(\{SS1,SS2,...,SS8\})$).

4.2 Reliability

Assuming that each HNS component may fail, we evaluate the system-wide reliability of HNS from a viewpoint of the availability of the integrated services. For a given HNS with scenarios, we define *n-reliability* as the probability that at least n scenarios are available in the HNS. The n-reliability varies depending on the architecture as well as the reliability of each component. Evaluating the reliability at the design process is crucial for reliable system implementation.

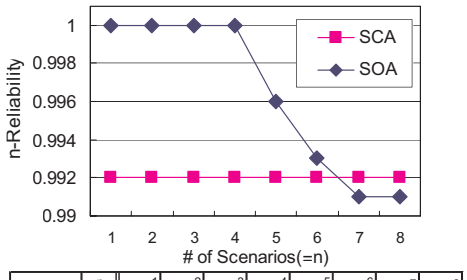
To calculate n-reliability, we apply the Sum of Disjoint Products (SDP) approach (Hariri, 1987; Soh, 1991; Tsuchiya, 2000) to the service integration graph. The SDP is a method to derive the network reliability based on pathset and cutset of the graph theory. Intuitively, when a graph G and reliability of each node (and edge) are given, the SDP method

calculates reliability that at least one of specified set of subgraphs of G is available (i.e., *operational*), by taking the overlaps among the subgraphs into account.

As seen in the previous subsection, each scenario in HNS is characterized by a *SIG*, and a *SIG* is a subgraph of *FSIG*. Hence, n-reliability can be calculated by SDP in such a way that some n *SIGs* are operational in *FSIG*. For instance, in our target HNS, 1-reliability is calculated by SDP as a probability at least one of *SIG*(SS1), ..., *SIG*(SS8) is operational. Similarly, 2-reliability is derived from *SIG*({SS1,SS2}),*SIG*({SS1,SS3}), ..., *SIG*({SS7,SS8}). Thus, taking all combinations from the given set of scenarios, we can compute n-reliability with the SDP method.

To evaluate the reliability purely relevant to the architecture of HNS, we assume that only WS (in SOA) and HS (in SCA) may fail. As an expected value, we set the reliability of each WS (in SOA) to be 0.999. We also set the reliability of the HS (in SCA) to be 0.992 (= 0.999^8, since HS implements proprietary programs for 8 scenarios). Then, we applied the SDP method to the SOA-based HNS (in Fig. 3) and the SCA-based HNS (in Figure 4).

The result is shown in Figure 5. The horizontal axis represents the number of scenarios (n), while the vertical axis plots n-reliability. From the result, it can be seen that n-reliability for SCA becomes equal to the reliability of HS. This is because all scenarios depend on the centralized HS. In other words, if the HS fails, all the scenarios become unavailable. On the other hand in SOA, the eight scenarios use distributed WS. Hence, even if a WS crashes, scenarios are partially operational. Thus, the SOA-based HNS achieves higher fault tolerance than the SCA-based HNS. For n = 7,8, SCA achieves slightly more reliable than SOA. This is because the probability that all the components in SOA are operational becomes smaller than that of SCA, since SOA contains more components.



	n	1	2	3	4	5	6	7	8
Architecture Type	SOA	0.99999	0.99999	0.99999	0.99998	0.996	0.99302	0.99104	0.99104
	SCA	0.992	0.992	0.992	0.992	0.992	0.992	0.992	0.992

Figure 5: Reliability.

Table 1: Workload.

(a)SOA-based HNS (b)SCA-based HNS

WS	WL(WS)
DVD_WS	10.7
TV_WS	29.8
Speaker_WS	29.8
Light_WS	57.4
Illumino_WS	57.4
Door_WS	18.7
Phone_WS	3.7
AC_WS	16
Thermo_WS	16
StandardDev	18.203

HomeServer	WL(HS)
HS	86.2
StandardDev	86.2

4.3 Workload

Our interest here is to measure a workload of each component (WS or HS) imposed when performing integrated services in HNS. The workload varies depending on the *usage frequency* of scenarios. Based on the given usage frequency, we characterize the workload of each component v as a total number of appearance of v in all scenarios. This metric enables us to determine the deviation of workload in HNS, so that we can change the design of HNS in consideration of load-balancing.

Suppose that we have $FSIG = (N, L, E)$ and scenarios $s_1, ..., s_n$. Also suppose that f_i $(1 \le i \le n)$ is a given usage frequency of scenario s_i. For each node $v \in N$, we define an *appearance function* $c_i: N \rightarrow \{0,1\}$ such that: $c_i(v) = 1$ iff v appears in $SIG(s_i)$, otherwise $c_i(v) = 0$. Then, a workload for the component v is defined by $WL(v) = \sum_{i=1}^{n} f_i \times c_i(v)$

For the evaluation of our target HNS, we interviewed 12 users (8 singles, 2 married men without children, 2 men with a family of four). We asked them the estimated usage frequency of the scenarios SS1 to SS8 per week, and obtained the average number of usage of each scenario. Based on this, we calculate the workloads of WS (in SOA) and HS (in SCA).

The result of the workload estimation is shown in Table 1. The column *WL*(WS) shows how many times each WS (or HS) is used per week. The result for SOA gives important information on which components require load-balancing. For example, since *WL*(Light_WS) is large, it would be reasonable to prepare a backup WS to share the load. Thus, in the SOA-based HNS, it is relatively easy to perform flexible design changes reflecting the workload. On the other hand, from the result of SCA, we can see that HS suffers from much heavier workload than those of SOA. The only way to perform the load-balancing is

duplicate the HS, which is not as flexible as the case of SOA.

4.4 Functional Complexity

In this subsection, we estimate the functional complexity for each component at the design stage. Basically, the functional complexity for a component v depends on how many methods v has to *provide* and *use*, in order to achieve all the integrated services. This is a key factor for implementing v. Specifically, for each node v in FSIG, we count the number of the labels attached to the incident edges of v.

Let $FSIG = (N, L, E)$ be given. An edge $(WS_A,m,WS_B)\in E$ describes that WS_A uses the method m of WS_B by definition. So, the function of m should be implemented inside WS_B. In this sense, we call m internal function of WS_B. On the other hand, from the viewpoint of WS_A, WS_A has to call m which is outside WS_A. Hence, m is called external function of WS_A.

For each component $v\in N$ in *FSIG*, we define the functional complexity of v as the number of internal and external functions of v. Strictly speaking, the number of internal functions of v is defined as $inum(v) = |\{m|\exists v';(v',m,v)\in E\}|$. Also, the number of external functions is defined as $enum(v) = |\{m|\exists v';(v,m,v')\in E\}|$. Then, the functional complexity of v is defined by $fcomp(v) = inum(v) + enum(v)$.

For example, let us take Light_WS in Figure 3. Then, *inum*(Light_WS) = |{1-1:Light.ON, 2-2:Light.ON, 3-7:Light.ON, 8-4:Light.OFF}| = 2 *enum*(Light_WS) = |{1-2:Illuminometer.ON, 1-3: Illuminometer.getIllumination, 2-3: Illuminometer.ON, 2-4: Illuminometer.getIllumination, 3-8: Illuminometer.ON, 3-9: Illuminometer.getIllumination, 8-5: Illuminometer.OFF, LightON, LightOFF, LightBrightnessControl}| = 6

Table 2 shows the functional complexity for all the components of our target HNS. It can be seen that each WS in SOA requires a smaller number of functions than HS in SCA. This implies that the effort taken for the implementation of WS would be smaller than that of SCA. Also for the SOA-based HNS, it is also possible for the designer to make the functional complexity well-balanced, by carefully modifying the scenario design (i.e., changing the topology of *FSIG*).

4.5 Coupling

The coupling measures the degree of dependence of a component against other components. Although the coupling between WS (in SOA) is basically loose (see Section 2), it provides a reasonable guideline for robust scenario designs. If a WS v is used by (or uses) a lot of other components, failure of v affects these components, which dramatically decreases availability of the service scenarios.

Table 2: Complexity and Coupling.

	Complexity		Coupling	
WS/HS	inum(WS)	enum(WS)	use(WS)	used(WS)
DVD_WS	2	6	3	1
TV_WS	5	7	2	3
Speaker_WS	4	4	1	1
Light_WS	2	6	2	3
Illumino_WS	3	3	1	1
Door_WS	1	4	3	1
Phone_WS	1	2	2	1
AC_WS	2	6	2	2
Thermo_WS	3	3	1	1
HS	8	23	9	1

Let $FSIG = (N, L, E)$ be given. For each component $v\in N$, we define *coupling* of v as the total number of components that v uses or are used by v. Strictly speaking, for $v\in N$, let $use(v) = |\{v'|\exists m;(v,m,v')\in E\}|$ and $used(v) = |\{v'|\exists m;(v',m,v)\in E\}|$. Then, coupling of v is defined by $coup(v) = use(v) + used(v)$.

For example, let us take TV_WS in Figure 3. Then, *use*(TV_WS) =|{speaker_WS, TV}|=2 *used*(TV_WS) =|{a user, DVD_WS, telephone_WS}|=3. Hence, *coup*(TV_WS) = 5.

The coupling for all the components of our target HNS is shown in right-half of Table 2. It can be seen in that the coupling of all WS (in SOA) is well-balanced. We can also see that the components in SCA are heavily dependent on the HS. This implies that the crash of HS is fatal, which is as discussed in Section 4.2.

5 DISCUSSION AND CONCLUDING REMARKS

In this paper, we proposed an application of the service-oriented architecture to HNS. We also presented a graph-based method to evaluate the architectural design of HNS. With a case study, we evaluated the HNS design using the four kinds of metrics and discussed the difference between SOA and SCA, quantitatively.

Of course, there are other important factors that we could not cover in this paper; such like performance, security, and implementation issues, etc. Therefore, we cannot say that the proposed SOA-based HNS is absolutely superior to the

conventional SCA-based HNS. Instead, our contribution is to show the applicability of SOA to the HNS through quantitative evaluation.

In the market of home electric appliances, a shift from dumb appliances to intelligent appliances is imminent. More convenient and more sophisticated integrated services will be required in the HNS such as entrance management and apparatus operation with user's voice. To make full use of such intelligent appliances in the HNS, SOA is quite promising architecture, as shown in this paper.

Another contribution is to present the concrete evaluation method for the architectural design of HNS. The proposed metrics provide useful information for the HNS developers to make various decisions on design, implementation and usability of HNS, at the early stage of the development.

Some topics for future research present themselves. We are currently implementing an HNS simulator with Web services in a distributed environment. The more practical evaluation using the simulator would allow us to find other useful metrics for the HNS development. In multi-user HNS, the *feature interaction* problem (Michael, 2003) must be considered, which is known as a functional conflict among scenarios and/or appliances. Investigating practical solution of the feature interaction problem is also our future work.

ACKNOWLEDGEMENTS

This work is partly supported by Grand-in-Aid for COE (Center Of Excellence) and Encouragement of Young Scientists (No.15700058), from Research of the Ministry of Education, Science, Sports and Culture, Japan.

REFERENCES

DHWG, (2004) Digital Home Working Group, [Online], Available: http://www.dhwg.org/ [2004].

ECHONET, (2004) ECHONET CONSORTIUM, [ONLINE],Available: http://www.echonet.gr.jp/ english /index.htm [2004].

Ethan, C., (2002) Web Services Essentials, United Stated of America: O'Reilly & Associates, Inc.

Hitachi. (2004) Horaso Network Service, [ONLINE], Available: http://ns.horaso.com/ [2004].

LG E., (2004) Home Network, [ONLINE], Available: http://www.lge.com/products/homenetwork/homenetwork.jsp [2004].

Michael, W., (2003) 'Feature Interactions in Web Services', Proc. of Seventh Int'l. Workshop on Feature Interactions in Telecommunication Networks and Distributed Systems (FIW'03), pp. 149-156.

Samsung. (2004) Home Network, [ONLINE], Available: http://www.samsung.com/HomeNetwork/index.htm [2004].

S. Hariri, and C. S. Raghavendra, (1987) 'SYREL: A Symbolic Reliability Algorithm Based on Path and Cutset Methods', IEEE Transactions on Computers, October, pp. 1224-1232.

Soh, S. and Rai, S., (1991) 'CAREL: Computer aided reliability evaluator for distributed computing networks', IEEE Trans. Parallel and Distributed Systems, July, pp. 199-213.

T., Tsuchiya, T., Kajikawa, and T., Kikuno, (2000) 'Parallelizing SDP (Sum of Disjoint Products) Algorithms for Fast Reliability Analysis', IEICE Transactions on Information and Systems, Vol. E83-D, No. 5, May , pp. 1183-1186.

W3C. (2004) W3C Web Service Activity, [ONLINE], Available: http://www.w3.org/2002/ws/ [13 Feb 2004].

Hao, H., (2003) What is Service-Oriented Architecture?, [ONLINE], Available: http://webservices.xml.com /pub/ a/ws/2003/09/30/soa.html [30 Sep 2003].

PART 2

Security and Reliability in Informations Systems and Networks

NEW NON-ADAPTIVE DISTRIBUTED SYSTEM-LEVEL DIAGNOSIS METHODS FOR COMPUTER NETWORKS

Hiroshi Masuyama
Information and Knowledge Engineering, Tottori University
Koyama-cho Minami 4-101, Tottori, 680-8552, Japan
Email: masuyama@ike.tottori-u.ac.jp

Koji Watanabe
Graduate School, Tottori University, Tottori, 680-8552, Japan

Keywords: Computer networks, System-level diagnosis, Diagnosability, Test graph.

Abstract: A hierarchical non-adaptive diagnosis algorithm is presented for testing total N nodes of computer networks. Since general computer networks can be regarded as an N-nodes complete graph, then for the efficient testing, it is essential that the test process be parallelized to enable simultaneous test of multiple nodes. In order to attain this object, we propose a noble test graph enabling to test as many nodes as possible in a network due to a hierarchical architecture of test processes. The amount of test times is evaluated as the diagnosis latency. Optimal diagnosability t is analyzed under clustered fault distribution. In order to reduce the amount of required test times, two revised approaches are discussed and evaluated.

1 INTRODUCTION

There have been significant theoretical researches in the area of **system-level** diagnosis by which every node receives diagnosis. This system-level diagnosis approach was introduced first by Preparata et al. (F. Preparata et al., 1968) where t-diagnosability was introduced. The **t-diagnosability** is the ability to diagnose a fault situation with t or fewer faults given in the network. This means that every node must be tested by more than t other nodes if a network is said to be t-diagnosable. The problems of fault detection (testing) and fault location (diagnosis) have been mostly studied by using testing networks which is reduced to some test graphs, whose vertices denote the nodes and whose an edge or test link p_i, p_j from node p_i to node p_j indicates that p_i tests p_j (C.Feng et al., 1996) ~ (N.H.Vaidya et al., 1994). Since a general graph contains many vertices, one by one test approach requires significant test time.

The fault model of the network characterizes the outcome of test results. The first model of system diagnosis is introduced as **PMC Model** (F. Preparata et al., 1968). In this model, the outcome of a test performed by a fault-free node is correct and equals fault state of the tested node. On the other hand, the outcome of a test performed by a faulty node is unreliable, that is, arbitrary. Classical system-level diagnosis approaches (F. Preparata et al., 1968), (S.L. Hakimi et al., 1974) have a central observer by which all test results are gathered to make a **syndrome** of the network. In the most of these approaches, a **distributed model** is assumed where each node performs independently its own local diagnosis, that is, performs tests of only its definite subset of nodes. If the choice of the next tests, that is, the subset is known in advance, these test approaches are also called a **non-adaptive** test. The central observer uses the results obtained from all test nodes to determine the fault situation, that is, locates the faults in the network.

On condition that a ring can be judged correctly whether the ring has at most one locatable fault or more than one un-locatable faults, a single loop testing (N.H. Vaidya et al., 1994) of one of **adaptive** diagnosis techniques where the choice of the next tests depends on the results of previous tests and not on a fixed pattern, is developed. There exist considerable presented schemes on the condition that the maximum number of faulty nodes distributed in a network is bounded by a predefined limit, and they have been improved to reduce the diagnosis latency (R.P. Bianchini et al., 1992), (E.P. Duarte Jr et al., 1998). However, since test graphs for general computer networks contains

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 117–124.

many vertices, these adaptive diagnosis techniques require significant overhead, that is complex analysis of the test results.

In this paper, we consider a classical system-level diagnosis algorithm in which only the nodes fail because a faulty communication link can be accommodated by treating as a faulty node. And we present a hierarchical non-adaptive diagnosis algorithm for testing total N nodes of computer networks. Since general computer networks can be regarded as an N-nodes complete graph, then for the efficient testing, it is essential that the test process be parallelized to enable simultaneous test of multiple nodes. In order to attain this object, we propose a regular graph of connectivity-$(t+1)$ with N nodes as test graphs. In this test graph, a self-tested node is placed at a key location in a hierarchical structure, and at first the node tests the adjacent nodes. Only adjacent nodes that passed the test can become new monitors and test their adjacent nodes, and so on. This process is propagated to higher levels of the test graph. At each level, all monitors send the announcements of their own test results "I passed the test" when they received a qualification as a monitor first, and in addition send only the test failed results of their test targets when they finish their tests, back to their monitors by which they are tested first. Each monitor also sends data transferred from its test target back to the monitor by which he is tested first. Then all test results are gathered in a host (that is, a central observer) directly connected the original monitor, and then the host can locate all faults in the network. Optimal diagnosability t is analyzed under clustered fault distribution.

Recently, several diagnosis techniques based on this self-testing (F.J. Meyer et al., 1989) have proposed, and achieved a successful diagnosis of a large number of faults. Though most of drawbacks of self-testing are to require many self-testing, papers (L. Zakrevski et al., 1998), and (H. Masuyama et al., 2001) made the drawbacks light by preparing the limited number of monitors, as shown in our approach. However, their target networks are multi-processor networks consisting of homogeneous nodes connected by bi-directional links. Each node can be viewed as a combination of a router and processor along with associated RAM, bus and I/O circuitry, then they differ from us in target networks.

In non-adaptive or even adaptive tests, since each node must performs a certain number of nodes and report to somewhere in the network, then a traffic problem must be cleared. Therefore, not only the time elapsed for testing all nodes and the time complexity of diagnosis algorithm but also the traffic condition are essential to evaluate diagnosis algorithms. In this paper, **diagnosis latency**, that is, the time elapsed for testing all nodes is evaluated as the total number of test times where each test executes in different time. This time is also called as testing round. In order to reduce the amount of required test times, two revised approaches are discussed and evaluated.

2 ALGORITHMS

In this section, we will discuss three algorithms for constructing our test graph, for obtaining necessary test orders, and for test.

2.1 Test Graph

For given N and diagnosability t, we will plan to construct a test graph whose connectivity is over t by the following algorithm:

[Algorithm A]

Step 1: Prepare α hypercubes of dimension β independently, and number to these α hypercubes. Each node in a hypercube corresponds to $(\alpha-1)$ nodes in each different hypercubes.

Step 2: For total β sets of α corresponding nodes, connect α corresponding nodes with a completed graph.

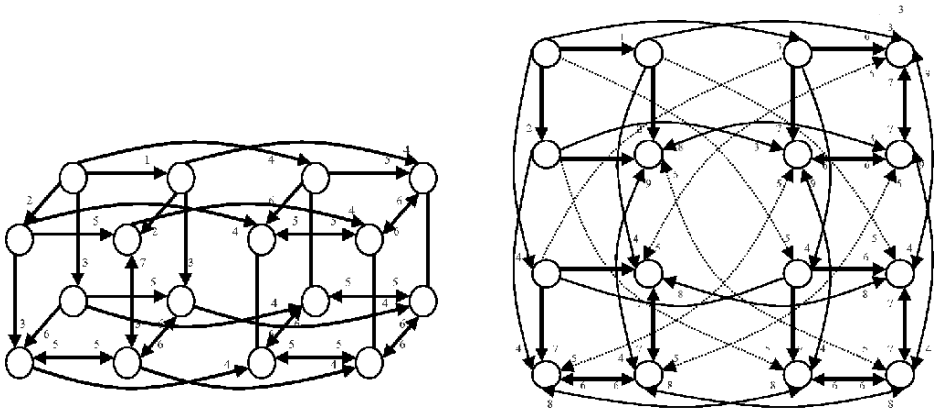
Step 3: Select one node as an original monitor arbitrary from N nodes. Set the edges connected with the original monitor and the adjacent nodes as unidirectional edges and all other edges as bidirectional edges.

The graph obtained by Algorithm A has $\alpha \cdot 2^{\beta}$ nodes, and the degree of each node is $\alpha + (\beta - 1)$. Then, α and β are restricted by given N and t as follows: $N = \alpha \cdot 2^{\beta}$ and $t \le \alpha + (\beta - 2)$. The longest distance d_m from an original monitor is $\beta + 1$.

On the strength of algorithm A for constructing test graph, we can give test orders to every adjacent nodes of each node by the following algorithm:

[Algorithm B]

Each node of a β-dimensional hypercube can be indexed 0 to $2^{\beta-1}$, and each of α hypercubes can be numbered 0 to $\alpha - 1$. Assume node i is indexed j and hypercube which contain node i is numbered $k(0 \le k \le \alpha - 1)$. The test orders of each adjacent node of node i are as follows: The adjacent nodes indexed $(j+1), (j+2), \cdots, (j-2), (j-1) \pmod{2^{\beta}}$ on hypercube numbered k, the the adjacent nodes on hypercubes numbered $(k+1), (k+2), \cdots, (k-2), (k-1) \pmod{\alpha}$.



(a) In the case of $\alpha = 1,\ \beta = 4,$ and $t = 2$ (b) In the case of $\alpha = 4,\ \beta = 2,$ and $t = 4$

Figure 1: Two test graphs with $N = 16$.

2.2 Test Algorithm

On the strength of Algorithms A and B, we can construct a test algorithm for an $N(=\alpha \cdot 2^{\beta})$-node network as follows:

[Algorithm C]

First, the monitor tests its adjacent nodes in the test order of the adjacent nodes, and hands a message "faulty node name" to the host if it decides an adjacent node faulty. The monitor hands a qualification as a monitor to its adjacent node if it decides the adjacent node non faulty.

Each node hands first its own test result "I passed the test" to its the first tester when it received a qualification as a monitor. Each node starts testing its adjacent nodes in the test order, and hands a message "faulty node name" to the adjacent node by which it is tested first if it decides its testing adjacent node faulty. It hands a qualification as a monitor to its adjacent node if it decides the adjacent node non faulty. Each node hands messages of "faulty node name" to the adjacent node by which it is tested first if it receives the messages from the adjacent node to which it tested previously.

Then, with Algorithm C all test results can be gathered in a host directly connected the original monitor, and then the host can locate all faults in the network.

Example 1: Figs. 1(a) and (b) show two test graphs with N=16 labeled the test orders in the cases of $(\alpha = 1,\ \beta = 4, t = 2)$ and $(\alpha = 4,\ \beta = 2, t = 4)$, respectively. Figs.2(a) and (b) show two test graphs with N=32 in the cases of $(\alpha = 2,\ \beta = 4, t = 4)$ and $(\alpha = 4,\ \beta = 3, t = 5)$, respectively.

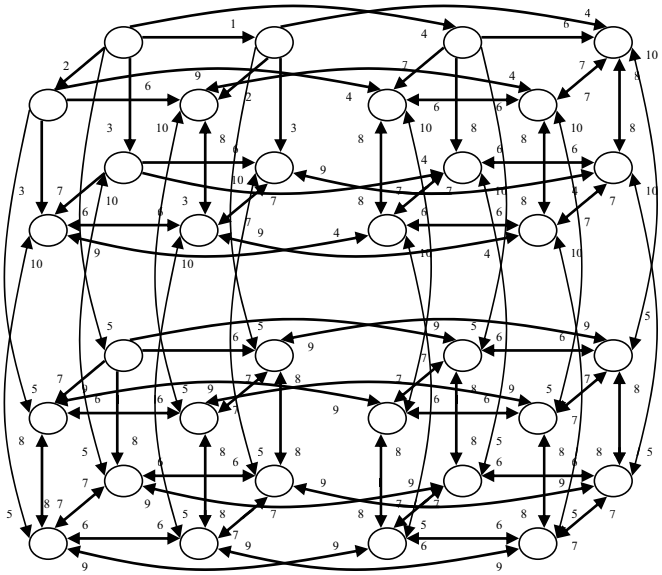
3 EVALUATION

3.1 Number of Test Times

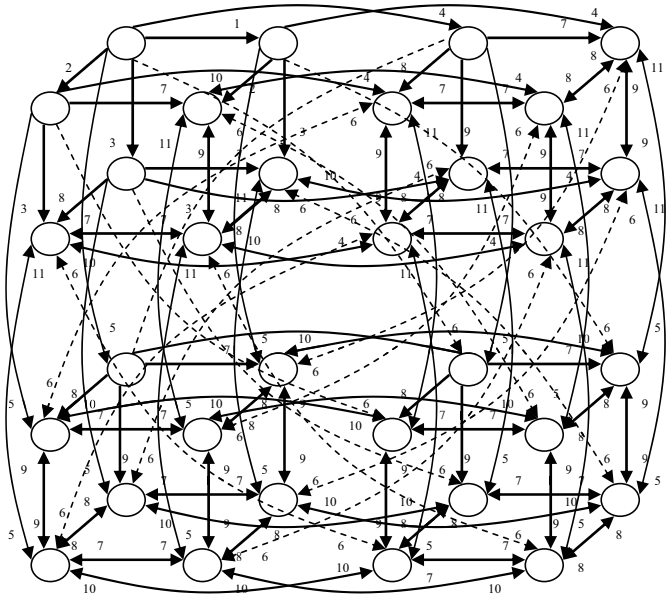
The total number of edges in a test graph with $N = \alpha \cdot 2^{\beta}$ and $t = \alpha + (\beta - 2)$ is $(N - (\alpha + \beta))(t + 1) + (\alpha + \beta - 1)$, where we count a bidirectional edges as 2 edges. This value becomes close to $N(t+1)$ when N is large. Let the total number of test times where each test executes in different time be T. Since the total number of nodes is N, then the number of tested arcs can increase exponentially up to N by taking test time γ which satisfies $N = 2^{\gamma}$. After the time γ, since the total number of tested arcs is $\sum_{i=0}^{\gamma-1} 2^i$,the number of untested arcs is $N(t+1) - \sum_{i=0}^{\gamma-1} 2^i$. These $N(t+1) - \sum_{i=0}^{\gamma-1} 2^i$ untested arcs can be tested N to N every test time, then it takes total $\left\{ N(t-1) - \sum_{i=0}^{\gamma-1} 2^i \right\} / N$ times.

Therefore, T is given as follows:

$$\begin{aligned} T &= \gamma + \left\{ N(t+1) - \sum_{i=0}^{\gamma-1} 2^i \right\} / N \\ &= \gamma + t \\ &= \log N + t \end{aligned} \quad (1)$$



(a) In the case of $\alpha = 2,\ \beta = 4,$ and $t = 4$.



(b) In the case of $\alpha = 4,\ \beta = 3,$ and $t = 5$

Figure 2: Two test graphs with $N = 32$.

Table 1: Probability of correct diagnosis for realistic yield in $N = 2^{13}$.

t	Yield (%)				
	99.999	99.750	99.500	99.250	99.000
1	1.0000	0.9616	0.8551	0.7031	0.5420
2	1.0000	0.9998	0.9989	0.9951	0.9915
3	1.0000	1.0000	1.0000	0.9999	0.9999
4	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000
6	1.0000	1.0000	1.0000	1.0000	1.0000

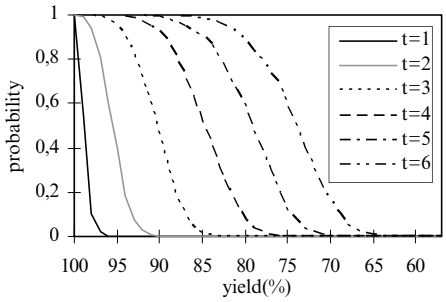


Figure 3: Probability of correct diagnosis for 6 diagnosabilities and $N = 2^{13}$.

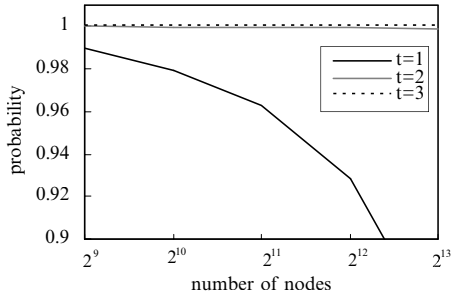


Figure 4: Probability of correct diagnosis for 5 network scales in $Y = 99.5\,\%$.

3.2 Time Complexity of Diagnosis Algorithm

Each node can test its adjacent nodes asynchronously in the test order which is given automatically by the test graph. Therefore, on the assumption that the time complexity of algorithm to test a node by the adjacent monitor is 1, the time complexity of diagnosis algorithm can be evaluated as the same as T.

3.3 Amount of Transmit Messages

Each node hands a message "faulty node name" to the adjacent node by which it is tested first if it decides its testing adjacent node faulty. Then, these messages "faulty node name" pass through at most $t(t+1)d_m$ edges in a test graph. The average amount of transmit messages on an edge is given as $t(t+1)d_m / N(t+1)$, that is td_m / N.

3.4 Analysis of Diagnosability T Under Clustered Fault Distribution

Extensive simulations were performed for evaluating the diagnosability when faulty nodes are clustered in a system. The examined systems consist of $2^{10} \sim 2^{13}$ nodes. A thousand different configurations of clustered faulty nodes in a system were simulated using negative binominal distributions. The diagnosis algorithm was run on all these configurations. Figure 3 gives the probability of correct diagnosis for the 6 scenarios of diagnosability and $N = 2^{13}$. It can be observed from Fig. 3 that, for any yield Y, the probability of correct diagnosis is higher for higher diagnosability. Thus, the diagnosis with t=1 has the least probability of correct diagnosis over all yields, as was expected. What we need to know is the smallest diagnosability by which diagnosis is correctly performed under the

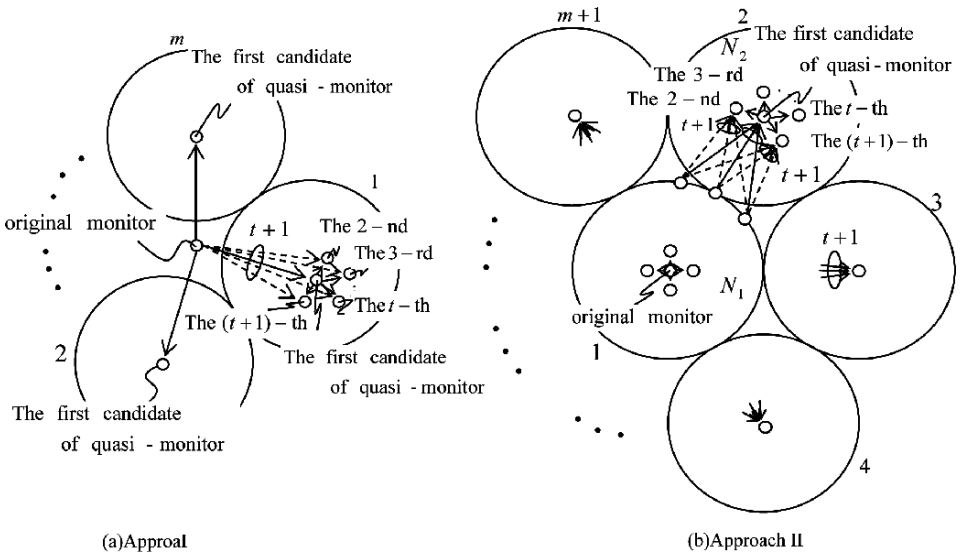


Figure 5: Two reduction approaches.

limits of realistic circumstances. Table 1 gives the probability within the realistic yield values in $N = 2^{13}$. Figure 4 gives the probability for the 5 scenarios of network scale in the case of Y = 99.5%. These data show an answer that t = 2 is proper.

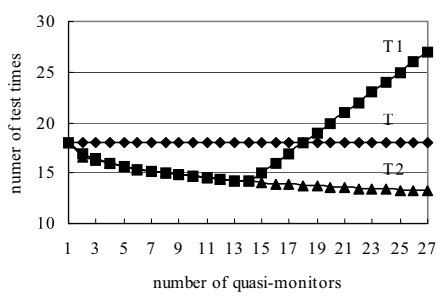
4 REDUCTION OF DIAGNOSIS PROCESS

In this section, we consider a technique to reduce the number of test times. Two approaches can be proposed as follows:

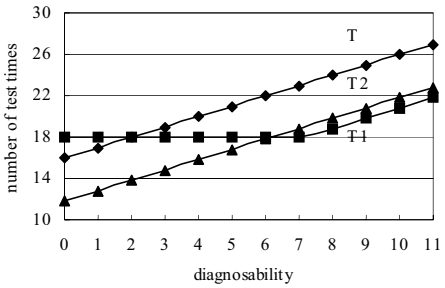
Let us set m quasi-monitors which perform the same test processes as the original monitor's. Since these quasi-monitors are not connected directly with the central observer, the gathered test results (faulty node names with its tester name) are stored temporarily in each quasi-monitor until each quasi-monitor receives a qualification as a monitor. After that, the quasi-monitor hands its test results to its own tester. The tester next hand the test result to the tester's tester, and so on. Finally, the test results is transmitted to the central observer. On this condition, we can consider two approaches to test the quasi-monitors as shown in Fig. 5. In Fig. 5(a), the original monitor tests only m quasi-monitors, then it does not test any other node. In Fig. 5(b), the original monitor does not test any quasi-monitor directly, then each quasi-monitor is tested by the adjacent nodes obtained a qualification as a monitor. This reformed point is that both original and quasi monitors enter for testing simultaneously. The un-inscribed part in each circle in Fig. 5 means the same structure as the test graph shown by Algorithm A. Each quasi-monitor hands its stored test results to its tester in order, as mentioned above. Then all test results can be gathered in a host directly connected the original monitor, and then the host can locate all faults in the network.

From the above discussion, we can understand the intention to reduce the number of test times, that is, the test graph can be partitioned into m (in Fig. 5(a)) or $m + 1$ (in Fig. 5(b)) parts by preparing m quasi-monitors. When the first candidate of quasi-monitor is judged as faulty, the second candidate is next tested, and so on. When an adjacent node of the first candidate of quasi-monitor is judged as non-faulty, the node takes the place of the first candidate of faulty quasi-monitor. The new quasi-monitor begins testing its adjacent nodes from the beginning.

Let us consider the relative merits of the above two approaches in the point of the number of required test times. Let T_1 and T_2 be the numbers of test times required, when all the first candidates of quasi-monitor are non faulty, in the approaches

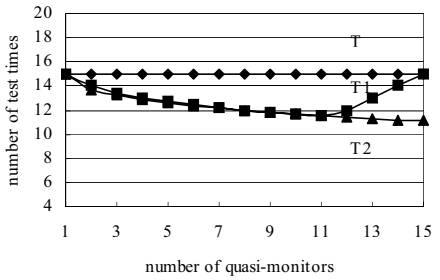


(a) In the case of $t = 2$

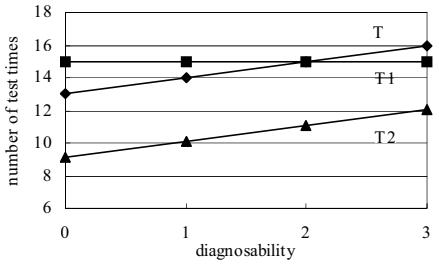


(b) In the case of $m = 18$

Figure 6: The relative merits in the case of $N = 2^{16}$.



(a) In the case of $t = 2$



(b) In the case of $m = 15$

Figure 7: The relative merits in the case of $N = 2^{13}$.

shown in Fig. 5 (a) and (b), respectively. That is, T_1 and T_2 are the smallest numbers of test times required in the approaches shown in Fig. 5 (a) and (b). We obtain the following two equations from eq.(1):

$$T_1 = \max\left[m, \log(N/m) + t - 1\right]$$

$$T_2 = \max\left[\log N_1 + t + 1, \log N_2 + t\right]$$

Were N_1 and N_2 are the total numbers of nodes in circles 1 and 2 in Fig. 5 (b), respectively. m is restricted by the following relationships:

$$(t+1)m \le N_1$$

$$N_1 + mN_2 = N$$

For simplification, we assume $mN_1 = N_2$, then we obtain T_2 and an inequality for m as

$$T_2 = \max\left[\log\left\{N/(m^2+1)\right\} + t + 1, \log\left\{mN/(m^2+1)\right\} + t\right],$$

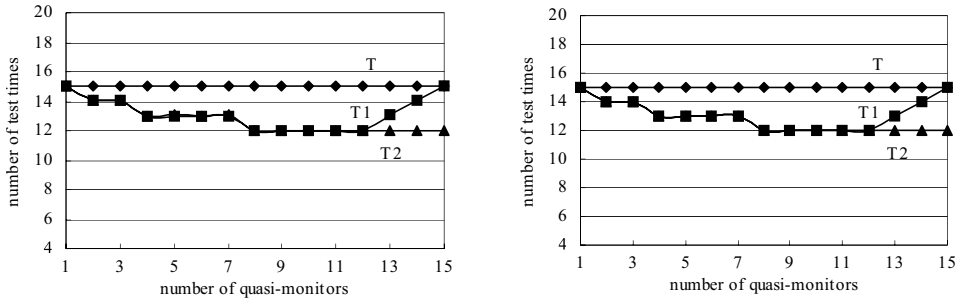
$$(t+1)(m^2+1)m \le N \qquad (2)$$

On the other hand, in the worst faulty case, that is, the biggest numbers $T_{1\max}$ and $T_{2\max}$ of test times required in Fig. 5 (a) and (b), respectively are as follows:

$$T_{1\max} \cong m + 2t + \log(N/m)$$

$$T_{2\max} \cong 3t + 2\log\left\{N/(m^2+1)\right\} + \log m$$

Figure 6(a) shows the relative merits of the above two and original approaches in the case of $N = 2^{16}$ and $t = 2$ under the restriction given by eq.(2). In this case, the boundary line of the relative merits is $m = 18$, that is, the scheme shown in Fig. 5(b) is superior to the others. On the other hand, Fig. 6(b) shows the merits in the case of $N = 2^{16}$ and $m = 18$ under the same restriction. In this case, the boundary line of the relative merits is $t = 6$, that is, the scheme shown in Fig. 5(a) is the best when t is over the boundary. Figure 7 shows relative merits in the case of $N = 2^{13}$, where the results show the same tendency as in the case of $N = 2^{16}$.



(a) In the case of $t = 2$, $N = 2^{13}$ and Y = 99.95% (b) In the case of $t = 2$, $N = 2^{13}$ and Y = 99.5%

Figure 8: The relative merits in realistic circumstances of fault pattern.

Extensive simulations were performed also for evaluating the relationship of the number of test times versus the number of quasi-monitors when faulty nodes are clustered in a system of 2^{13} nodes. A thousand different configuration of clustered faulty nodes in the system were simulated using negative binominal distributions on condition of t = 2. Figures 8(a) and (b) show the results in the cases of Y = 99.95% and 99.50%, respectively, where Y is the yield of nodes in the system. The same property as mentioned above is proved in realistic circumstances.

5 CONCLUSION

A hierarchical non-adaptive diagnosis algorithm is presented for testing total N nodes of computer networks. We proposed a noble test graph with $(t+1)$-connectivity enabling to test as many nodes as possible in a network due to a hierarchical architecture of test processes. If the maximum number of faulty nodes distributed in a network is bounded by a predefined limit t, our approach is effective. In this approach, an original monitor is placed at a key location in a network, and at first the monitor tests the adjacent nodes. Only adjacent nodes that passed the test can become new monitors and test their adjacent nodes, and so on. This process is propagated to higher levels of the test graph. At each level, every new monitor sends their information as a successful candidate (new monitor) back to a central observer directly connected original monitor through only one route. Monitor sends its test result back to a central observer through only one route if it decides its adjacent node faulty. Consequently, the observer can gather all information of faults in the network. The amount of test times is evaluated as the diagnosis latency. Optimal diagnosability t is analyzed under clustered fault distribution. Two revised approaches to reduce the required test times are discussed and the relative merits of three approaches are evaluated.

REFERENCES

F. Preparata, G. Metze, and R.T. Chien, “On the Connection Assignment Problem of Diagnosable Systems,” IEEE Trans. Electronic Computers, vol. 16, pp. 848-854, 1968.

C. Feng, L.N. Bhuyan, and F. Lombardi, “Adaptive System-Level Diagnosis for Hypercube Multi-Processors,” IEEE Trans. on Computers, vol. 45, no. 10, pp. 1157-1170, 1996.

C.R. Kime, “System Diagnosiss,” In Fault-Tolerant Computing: Theory and Techniques, vol. 2, D.K. Pradhan(ed.), Prentice-Hall, New Jersey, 1986.

D.P. Siewiorek and R.S.Swarz, “Reliable Computer System – Design and Evaluation,” 2nd ed. Digital Press, Bredford, MA, 1992.

N.H. Vaidya and D.K. Pradham, “Safe System Level Diagnosis,” IEEE Trans. Comput. Vol. 43, no. 3, pp. 367-370, 1994.

S.L. Hakimi and A.T. Amin, “Characterization of Connection Assignment of Diagnosable Systems,” IEEE Trans. Comput., no. 1, vol. C-23, 1974.

R.P. Bianchini and R. Buskens, “Implementation of On-Line distributed System-Level Diagnosis Theory,” IEEE Trans. Comput. vol. 41, no. 5, pp. 616-626, 1992.

E.P. Duarte Jr. and T. Nanya, “A Hierarchical Adaptive Distributed System-Level Diagnosis Algorithm,” IEEE Trans. Comput. Vol. 47, no. 1, pp. 34-45, 1998.

F.J. Meyer and D.H. Pradhan,”Dynamic Testing strategy for Distributed Systems,” IEEE Trans. Comput., vol. 39, no. 3, pp. 356-365, 1989.

L. Zakrevski and M.G. Karpovsky, ”Fault-Tolerant Message Routing for Multiprocessors.” Parallel and Distributed Proscessing (Edited J.Rolim), Springer, pp. 714-731, 1998.

H. Masuyama, Y. Ohashi, and T. Miyoshi, “A Diagnosis Method of Computer Networks.” 2001 Proceedings of IASTED Parallel and Distributed Computing and Systems, pp. 474-479, 2001.

GSM AND GPRS PERFORMANCE OF IPSEC DATA COMMUNICATION

Gianluigi Me, Giuseppe F. Italiano
Dipartimento di Informatica, Sistemi e Produzione, Roma, Italy
Email: {me,italiano}@disp.uniroma2.it

Paolo Spagnoletti
CeRSI, LUISS "Guido Carli" University,via Tommasini 1, Roma, Italy
Email: pspagnoletti@luiss.it

Keywords: Mobile application, Security.

Abstract: Cellular Internet services must grapple with the added security threats posed by the radio transmission, open to eavesdropping. Furthermore, the combination of always-on connectivity and an interface to the public Internet means high speed data services has to cope with the same security issues that can be found in the wired environment.Confidentiality of GSM/GPRS communications has been provided only in BS-ME/GGSN-ME by COMP128/GEA+ algorithms, whose strength is often not believed adequate for corporate/governmental requirements. Furthermore, A5/1 and A5/2 algorithms have been recently attacked with real time ciphertext only cryptanalysis by Barkan, Biham and Keller. To provide an adequate level of security, it is often argued to employ IPSec over the GSM/GPRS framework. We provide experimental evidences that IPSec is a viable solution to provide the desired level of security. In particular, the overhead generated is tolerable where high sensitive/critical communications take place. We expect that our findings could help better understanding how securing a deployed GSM/GPRS network which corporate/governmental infrastructures can rely on and what performances can be expected by using IPsec over these media.

1 INTRODUCTION

Wireless technology is widespread in today's communication networks, mainly due to its facility of deployment and management. However, many security concerns about wireless infrastructures have been raised in recent years. In particular, there has been a serious consciousness of the weaknesses of GSM and GPRS, among other wireless technologies. Important works on this area are by Barkan et al. 0, Biryukov et al. 0, Briceno et al. 0 and Ekhdal et al. 0, whose pose serious threats to GSM/GPRS, with high-cost/easy to use systems. Furthermore GSM is the most widely used cellular technology, with more than 787.5 million customers in over 191 countries.

All these facts make these two technologies highly insecure and untrustable for who has to communicate with confidentiality and suggested us to propose a secure architecture for people/corporate/government with security requirements. In this paper, we analyze the overhead introduced to secure GSM/GPRS communication. In particular, we investigate the performance of the IPSec protocol employed to secure communication over GSM/GPRS. We show with experimental results that for a wide range of parameters, the overhead introduced by the IPSec is limited. Hence, we experimentally argue that the adoption of the IPSec suite is a viable solution to secure public GPRS network infrastructure.

The remainder of the paper is organized as follows: firstly a security background, where we briefly highlight the security features and the threats to which the GSM/GPRS is subject to. Then, we detail our security architecture implementation, focusing on relevant IPSEC countermeasures to GSM/GPRS threats. Finally we develop our

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 125–133.

consideration on IPSec encryption over GSM/GPRS. In particular, we will illustrate the methodology adopted to perform the measurement and the result of our analysis, based on a wide range of experiment that have been carried out, varying different, sensitive parameters of interest of the IPSec suite and the type of traffic secured.

2 GSM/GPRS STANDARD SECURITY

The Global System for Mobile Communications (GSM) (Figure 1) security was designed with three constraints in mind 0: α) Concern of granting too much security and so bringing export problems upon GSM; β) GSM did not have to be resistant to active attacks where the attacker interferes with the operation of the system, perhaps masquerading as a system entity; and γ) The trust between operators for the security operation should be minimized. The use of air interface at the transmission media allows a number of potential threats from eavesdropping. As stated by 0, it was soon apparent in the threat analysis that the weakest part of the system was the radio path, as this can be easily intercepted. In fact, there was no attempt to provide security on the fixed network part of GSM.

The General Packet Radio Service (GPRS) is a GSM-based service which provides mobile users with true packet access to data network. GPRS uses a packet-mode technique to transfer high-speed and low-speed data and signaling in an efficient manner 0. Security in GPRS is largely based on the GSM system security function. The main entities involved are the SGSN (Serving GPRS Support Node), GGSN (Gateway GPRS Support Node), AuC (Authentication Center) and HLR (Home Location Register). The HLR and AuC provide the same functionality as in GSM. The SGSN and GGSN both take care of authentication (Figure 1). The main functions related to GPRS device (MS) are authentication and encryption.

The authentication in GSM systems happens in VLR (Visitor Location Register) or HLR 0, through an Authentication Key (Ki) 0 stored in the AuC of the home PLMN (Public Land Mobile Network), using A3 (0, 0) algorithm. The operators may be free to design their own A3 algorithm.

Nearly every GSM operator in the world uses an algorithm called COMP128 for both A3 (authentication) and A8 (key generation) algorithms 0. The GPRS authentication procedure is handled in the same way as in GSM with the distinction that the procedures are executed in the SGSN. In some cases, the SGSN requests the pairs for a MS from the HLR/AuC corresponding to the IMSI of the MS. The GSM voice calls are encrypted using a family of algorithms collectively called A5. A5/0 uses no encryption. A5/1 is the "standard"export limited encryption algorithm,while A5/2 is the "export" (weakened) algorithm. A5/3 is a new algorithm based on the UMTS/WCDMA algorithm Kasumi 0. In GPRS network the ciphering scope is different: in GSM the scope is between BTS (Base Transceiver Station) and MS, in GPRS the scope is from the SGSN to the MS. The GPRS ciphering, performed at the LLC layer, is done with a family of algorithms: GEA0 (none), GEA1 (export), GEA2 (normal strength) and GEA3 (new, and effectively the same as A5/3).

2.1 GSM/GPRS authentication algorithms vulnerabilities

The protocol is simple, however, there are some vulnerabilities posed by its use. Namely, the TMSIs (Temporary Mobile Subscriber Identity) are generated based on the previous TMSI, therefore a missed synchronization in the TMSIs may require the IMSI to be used to set up it again, wherein the IMSI is sent in plaintext to the VLR, exposing its true identity. Also, there is no mechanism to prevent reply attacks. Once the session key Kc is compromised, by playing back the RAND, and the SRES, an intruder can impersonate the VLR since the protocol does not support network authentication.

Furthermore,

- Wagner and Goldberg announced in April 1998 that they had cracked COMP128 who had a weakness which would allow complete knowledge of Ki if around 160000 chosen RAND-SRES pairs could be collected (chosen plaintext attack). There are active attacks that can be used to obtain these pairs.
- The quickest attack would be to steal the user's mobile phone, remove the SIM and connect it to a phone emulator that can be used to send 160 000 chosen RAND to the SIM and receive the SRES. SIM tend to have relatively slow clock speeds and it

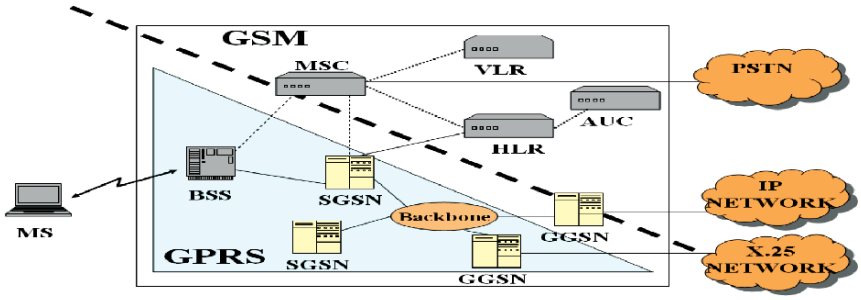


Figure 1

can therefore take up to 10 hours to obtain the 160000 pairs (with faster SIM, it would take 2 and a half hours).

- Retrieving the key from the SIM: the security of the whole GSM/GPRS security model is based on the secret Ki. If this key is compromised the whole account is compromised. Once the attacker is able to retrieve the Ki, he can not only listen to the subscribers calls, but also place calls billed to the original subscriber's account, because he can now impersonate the legitimate subscriber.
- Another method is to perform man in the middle attacks. Using a false BTS to send the RAND over the air interface, the rate at which pairs can be collected is slower and would take a number of days; however the attacker does not need physical possession of the SIM. After these efforts, the attacker has the Ki and can masquerade as the user and run calls on his bill, and also determine the Kc for the user's calls and therefore eavesdrop upon them 0;
- Cloning attack to A3 is further presented in 0.

The following attacks represent threats for authentication in GPRS:

- Spoofed Create PDP (Packet Data Protocol) Context Request: GTP (GPRS Tunnelling Protocol) inherently provides no authentication for the SGSNs and GGSNs themselves. This means that given the appropriate information of a subscriber, an attacker with access to the GRX (GPRS Roaming Exchange), another operator attached to the GRX, or a malicious insider can potentially create their own bogus SGSN and create a GTP tunnel to the GGSN of a subscriber. They can then pretend to be the legitimate subscriber when they are not. This can result in an operator providing illegitimate Internet access or possibly unauthorized access to the network of a corporate customer;
- Spoofed Update PDP Context Request: An attacker can use their own SGSN or a compromised SGSN to send an Update PDP Context Request to an SGSN, which is handling an existing GTP session. The attacker can then insert their own SGSN into the GTP session and hijack the data connection of the subscriber.

2.2 GSM/GPRS confidentiality algorithms vulnerabilities

The confidentiality of the GSM architecture is not completely sound. In the following we highlight a few security flows that have been published in literature. Our aim is not to discuss the GSM architecture nor its cryptographic flaws, but only showing that the native confidentiality it provides is weak, thus justifying the adoption on another independent security layer, as IPSec is. Furthermore, the security of the GSM confidentiality is based on the security through obscurity paradigm, debatable choice and usually leads, sooner or later, to system compromising 0.In the following paragraphs, we overview the main known attacks, paying the best attention to 0:

- Brute-force attack against A5. A real-time brute-force attack against the GSM security system is not feasible, since the time complexity is far too big, but with the distributed computer systems we can drastically reduce the time required;
- Divide-and-conquer attack against A5 – a divide-and-conquer attack is based on a known-plain-text

attack and can dramatically reduce the complexity (up to 2^9 – 2^14) 0;

- The only attack on an algorithm that has been confirmed to be A5/1 was that by Biryukov and Shamir, later improved by Wagner. The technique used is known as is time-memory trade off 0;
- Accessing the operator's signaling network: the airwaves between the MS and the BTS are not the only vulnerable point in the GSM system. The transmissions are encrypted only between the MS and the BTS. After the BTS, the traffic is transmitted in plain text within the operator's network. If the attacker can access the operator's signaling network, he will be able to listen to everything that is transmitted, including the actual phone call as well as the RAND, SRES and Kc;
- Real time cryptanalysis: the very new result, faced by our proposal architecture for data communi cation, comes from 0. The coding introduces known linear relationships between the bits to be encrypted; so even though the attacker might not know the value sof particular input bits, they know that certain groups of them XOR to 0. So, taking the same groups of encrypted bits and XORing them reveals the corresponding XOR of the keystream bits. This is the fundamental problem that allows the attacks to work without any knowledge at all of what is being encrypted, which is what they mean by "ciphertext only". The important thing about the active attacks is that the attacker can confuse a mobile into doing what it wants the mobile to do. At the limit, if the atta cker has intercepted the random challenge sent to a particular mobile and has recorded all the traffic, whether it is GSM voice or GPRS data, they can later send the same random challenge to the mobile and tell it to use A5/2 to communicate. When the mobile responds, they recover the key, and it's the same key that will decrypt the recorded stuff, whatever it was encrypted with.

2.3 How IPSec Matches Security Requirements

In previous paragraphs we have shown the cryptographic vulnerabilities of GSM/GPRS. In this mobile environment we have identified the following requirements, not appropriately covered by GSM/GPRS: (Ra) Protecting sensitive information: assuring the confidentiality and integrity of communications; (Rb) Access Control and Authorization; (Rc) Upper IP layer system availability, to guarantee the best communication media DoS robusteness. Furthermore, we intend to address these specific threats considering that, in the GSM/ GPRS framework, performing traffic analysis pose more concerns due to the fact that digital IP based traffic carries source and destination IP addresses in cleartext.

Furthermore, α) this system doesn't face communication parties localization tracing problem, because inherently coupled with GSM/GPRS link layer; β) DoS attacks to GSM/GPRS link layer are out of the requirements scope of this paper.

Our IPSec based architecture matches these requirements as follows:

Ra) Confidentiality of 3DES, the algorithm used in his architecture, is definitively better than A5. Furthermore it's possible to choose the preferred encryption algorithm in the IPSec suite, e.g. AES. Integrity is performed by HMAC-MD5 (keyed hash) function.

Rb) Authentication is performed combining IPSec preshared-keys (device authentication) and One Time Password (user authentication). This further layer has been needed because preshared key authentication creates a master key that is less secure because of absence of Perfect Forward Secrecy.

Rc) This requirement is matched by using IKE (Internet Key Exchange) in main mode, not aggressive 0.

3 ARCHITECTURAL OVERVIEW

As shown in Figure 2, we used a laptop connected to a Merlin 3+1 GPRS phone (3 downlink, 1 uplink channels) through a serial PPP (point-to point) link to act as a GPRS mobile terminal. We tested this architecture in an operational environment, with an Italian mobile carrier. The firewall acts as VPN concentrator in the architecture, thus establishing an IPSec tunnel (end to end) between the mobile terminal and the firewall inside the laboratory LAN. The sniffer has been placed on a switch connected to the firewall external interface, the firewall internal interface, the authentication and the application server and the router connected by a 2 Mbps E1 with the carrier, observing all the packets exchanged between nodes of the architecture.

In the service provider's backbone network the support of GPRS is done adding two new network elements: the Serving GPRS Support Node (SGSN) and the Gateway GPRS Support Node (GGSN).

3.1 Security General Overview

The information exchange has been shown in Figure 2. User, after providing three usual pieces of information (User Name, Password, APN) to log into mobile carrier GPRS networks, starts the IPSec tunnel setup phase with the system. An encrypted tunnel mode is adopted, where the IP information and the data are encrypted with a new IP address created and mapped to the IPSEC endpoints. This solution provides the overall highest data privacy 0.

After IPSEC device authentication and encrypted channel establishment, user authentication follows, to guarantee the identity of the person using the IPSEC node. This is because an encrypted session is established between the two devices in different locations. The user authentication mechanism gives the access to origin server application, thus preventing the attacker from accessing the system just stealing the mobile device. The system presented in this paper adopts an authentication schema based on strong two factor, token based schema, requiring two elements to verify an user identity: a physical element in user possession (a hardware keyfob) and a code that only the token owner knows (PIN code).

Furthermore, the static IP address adopted enables a greater level of security on the VPN, since the server can recognize the IP addresses of the clients. A device attempting to connect with an IP address unrecognized by the server would be denied access. NAT, economical further security level, seems aviable solution to the limited number of IP addresses available, by allowing the use of an unregistered IP addresses within the organization.

3.2 Architecture and Set-Up

Our architectural framework is synthetically detailed in Figure 2. It encompasses the following components:

-*Wireless mobile client*; provided with a COTS wireless mobile laptop running an application with transaction features (BITS IPSec-Telnet over GPRS capability) that provide the set up of an IPSEC ESP tunnel, strong-encrypted user authentication, host access via Telnet capabilities.

-*Firewall/Proxy*, adopting the following standards: IETF IPSec Standard, IETF IKE Standard (ISAKMP/OAKLEY) and NAT. The following services are thus available: Security Association and Key Manager, Policy Storage Service Provider, Policy Relay Service Provider, Internet Key Exchange Service Provider, and IPSec Engine. The authentication relies on the Diffie-Hellman algorithm, used with "pre-shared" keys, Diffie Hellman Group Oakley Default Group 1. The negotiation algorithm is DES-CBC with an explicit Initialization Vector 0 with authenticator HMAC-MD5-96 0;

-*Authentication server*; authenticates the users requiring to connect to the Host gateway;

-*Host gateway*, provides the results of the query to the Host, where application data resides, in a Telnet format;

For these measurements the MTU was set to 1500. When the connection is established, each end set the MSS to 1460 bytes with a window size of 16384 bytes.

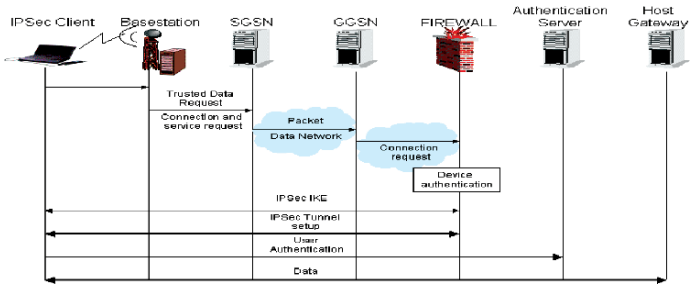


Figure 2

4 MEASUREMENTS METHODOLOGY AND SETTINGS

We analysed general statistics of the transformation made by IPSec, focusing on datagram sizes, basic step for an eavesdropper trying perform statistical cryptanalysis. The following considerations define the test environment:
- Bandwidth variation, changes of the bandwidth available for a connection throughout its lifetime, can represent a major acute problem. In fact, a number of factors may cause the connection's bandwidth variation. Change in the number or activity volume of other connections sharing the same bandwidth resource (e.g. the same time slot/s in GPRS networks), the narrowing/widening of the total bandwidth dedicated to data users (e.g. the start/end of voice calls in GPRS networks), and radio-link optimisations due to SNR changes are significant factors in bandwidth breathing. Failing to properly respond to those changes will result in the transport protocol either under utilizing the scarce wireless bandwidth or overflowing the network. A possible solution for these problems is presented in 0. Because of bandwidth variation our analysis is performed in the same low-traffic hours.
- The reliability in stationary connections is adequate, but the reliability in moving connections, with the same parameters is very poor. Therefore, the reliability of moving connections may create huge problems, if a distributed application cannot cope properly with disconnections or long pauses. This problem hardly relies on GSM/GPRS mobile operator capabilities. For this reason, presented measurements were performed with good to excellent signal coverage, since this threshold is the lowest boundary to enable the transaction, as we further investigate in next paragraph. An isolated, fixed test site was **set-up** to minimize influence from competing Internet traffic taking into account the needs of detailed measurements within lower protocol layers.

In general, in good radio signal quality environment, GPRS provides satisfactory throughput 0. The throughput and round-trip time in stationary connections were stable.

With respect of presented test environment fixed conditions, the analysis has been performed just once.

4.1 Methodology

Basing on GPRS network performance, we are interested in examining the performance of IPSEC over GPRS, evaluating performance and security strength and weaknesses of this solution, inspecting only the Ethernet traffic from two observation points located at the two sides of the firewall (encrypted and clear text).

The measurements refer to entire IP datagram length from LAN and GPRS side: in this architecture, we remark that IPSec works only on LAN IP payload, the LAN IP header is discarded and substituted with the firewall IP header. We did not perform any measurement on air link and we did not change TCP parameters (e.g. RTO, MSS, Congestion Window, SACK) during our measurements. After a general overview of traffic, we isolated traffic, keyed by state, on different channels (GPRS up/down link, LAN up/down link).

The keyed states refer to:

IKE exchange: directly inspectable by sniffing. Here, the main mode accomplish the establishment of ISAKMP SA, performed by IKE and DOI: a secure and authenticated communication channel (IKE SA) and authenticated keys used to provide confidentiality, message integrity, and message source authentication to the IKE communications between UDP exchanging packets on well-known port 500;

Device authentication: Then IPSec SA are established, and other protocol SAs can be negotiated; this phase starts on first ESP packet exchanged and we assume that finishes when the last but one ESP packet before we inspect on LAN traffic the first user authentication string. This assumption is correct as long as no Firewall - Authentication server interaction acts before the last but one ESP packet.

User authentication starts at next packet and finishes when the firewall delivers to the mobile user the ESP packets carrying the initial application form provided by the Host gateway. This form**,** triggering the application query**,** certifies that the user has been authenticated;

Transaction starts at next packet until the end;
Then we mapped this traffic segments on 4 different channels: the GPRS and LAN uplink and downlink as stated in Figure 3.

4.2 Measurements

The goal of our analysis was to compare protocol efficiency in data transfer using a telnet session encrypted with IPSEC on a GSM and a GPRS

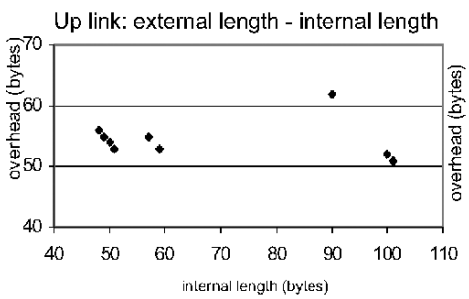


Figure 4a

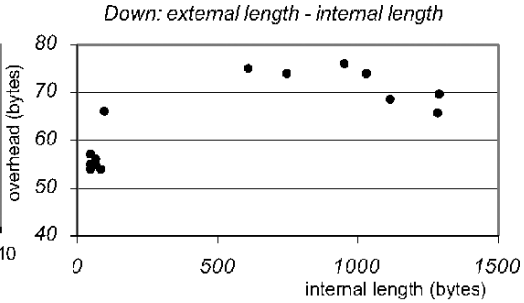


Figure 4b

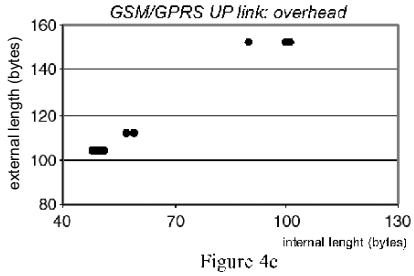


Figure 4c

channel in terms of:
- overhead and datagram fragmentation;
- time and costs.

Correspondently to all phases of session flow, IP datagram length matches in GSM/GPRS. This first straightforward result confirms the best expected forecast, due to independence of IP layer to the LLC (apart from some spurious "IKE INFORMATIONAL "datagrams). In fact, no re-transmission happened.

We made two complete connection from a MS using GSM and GPRS at the physical layer and we analysed IP datagrams exchanged between MT and External Firewall (EF) and between Internal Firewall (IF) and the Authentication Server (AS) dividing a complete transaction in three phases:
- device authentication (up-link and down-link);
- user authentication (up-link and down-link);
- telnet transaction (up-link and down-link).

The third phase was performed by executing a macro to eliminate the man latency in the editing phase of fields.

4.3 Overhead and Fragmentation

The overhead analysis demonstrates that:
- the maximum length of an IP datagram in the wireless path is 608 bytes;
- the maximum length of an IP datagram in the LAN path is 1061 bytes, due to the MTU of internal network;
- the overhead change with the length of datagrams;
- the behaviour in the GSM and GPRS case is exactly the same.

During the *transaction* phase the host gateway sends clear text packets to the internal firewall which performs the encryption retransmitting the packet over the wireless path to the mobile device (LAN and wireless down link). The encrypted packets leaving the firewall present an overhead due to the application of cryptographic algorithm performed by IPSec. The inverse happens to the encrypted packets transmitted by the mobile device.

We measured the discussed overhead in the two cases: up and down link. Because of difference between the MTU of LAN and GSM/GPRS paths, the first case (up link) is more simple than the down link case. In fact, the firewall receive, from the wireless link, always datagrams smaller than 608 bytes and after the decryption, it forwards clear text datagrams to the host gateway. In this case the IPSec overhead is represented on Figure 4a where Y axis measure the overhead corresponding to the internal datagram length specified on the X axis. The overhead range is 50-62 bytes, we will discuss later about the function linking overhead and internal length. In the down link case, for the fragmented datagrams (internal length > 608 bytes), we define the average overhead:

$(\sum_{0}^{N} L - \sum_{0}^{M} l) / N$

L= sum of fragmented datagrams length
l= sum of internal datagram length
N= number of external fragment

Figure 4b, shows how the average overhead is in the range 50-60 bytes for small datagrams (less than 608 bytes) and about 70 bytes for larger datagrams. Moreover, also in this case, there is an overhead variation for different values of internal datagram length and there are no significant differences between the GSM and the GPRS case.

To understand the relationship between overhead and datagram length we can observe from a different point of view what happens in the up link case. Figure 4c shows that the length of encrypted datagrams belongs to a discrete set of values. In particular, as the internal packet length increases, the length of external datagram assumes discrete set of increasing values. The reason of this behavior is the padding introduced by the encryption algorithm, useful to obfuscate statistical cryptanalysis.

How stated in 0, padding in an ESP packet is optional and the sender may add 0-255 bytes of padding. Padding is required when an encryption algorithm is employed that requires the plaintext to be a multiple of some number of bytes, or, irrespective of encryption algorithm requirements, to ensure that the resulting ciphertext terminates on a 4-byte boundary. Padding may be used to conceal the actual length of the payload, in support of (partial) traffic flow confidentiality. In this case, the inclusion of such additional padding has adverse bandwidth implications.

4.4 Time and Costs

We have already introduced some aspect about the time analysis and the difficult in performing a valid set of tests to compare the performances of GSM and GPRS links. In fact the bandwidth variation, the signal strength and the number of users simultaneously connected, made the transmission rate of GPRS variable between 0 and the maximum rate. Moreover, the performances of interactive traffic in the particular case of the link configuration phase of PPP increase the latency slowing the first phase of a GSM connection 0. With the performed analysis we have focused only on datagram length measurement to be sure that the results are independent from the factors discussed above. Moreover also in the presented case we observed that the GPRS was faster than GSM a part a delay in the “authentication device” phase, due to an IKE informational packet present in the GPRS case. The overhead introduced by encryption afflicts costs, with respect to bytes exchanged (GPRS) and connection time (GSM) of session flow. In fact, the above measurement shows that the overhead, varying in the 50-80 bytes range for each datagram, afflicts the traffic as follows:

- up link case: datagrams, containing mainly queries data, are doubled (small packets not longer than 70 bytes);
- down link case: datagrams containing application layer responses fragments (3270 format), are increased of 7-12% (datagram longer than 600 bytes).

We argue an average increment of traffic and costs, in the GPRS case, approximately of 10%.

Further studies can take into account GPRS bandwidth variation and the relationship with IPSEC performance in term of time and cost, with different session application (e.g. FTP, HTTP) and authentication and encryption protocols.

5 CONCLUSIONS

In this paper we have showed how the IPSec suite can be effectively applied to secure GSM/GPRS communications. The level of reliability in GSM/GPRS communications that this result can induce the deployment of large scale GPRS networks, as well as the adoption of public network GPRS-based, in critical governmental/private infrastructure. In particular, we have showed the effectiveness of the IPSec, proving that the overhead generated is tolerable under a wide set of parameters. The only limitation, posed by mobile operator capabilities, relies on GPRS connection reliability while roaming.

As for further research directions, we are interested in techniques to reduce the burst overhead generated by the set up IPSec-secured GPRS communications and to further study IPSEC connection reliability while roaming in GPRS environment. Moreover, we are addressing the possibility to employ the IPSec suite to secure peer to peer, ad hoc networks.

REFERENCES

Barkan, Biham and Keller, “*Instant Ciphertext-Only Cryptanalysis of GSM Encrypted Communication*”, Proceedings Crypto 2003 http://www.cs.technion.ac.il/users/wwwb/cgi-bin/tr-get.cgi/2003/CS/CS-2003-05.ps.gz, 2003.

Biryukov A, Shamir A, Wagner D., *"Real time cryptanaly sis of A5/1 on a P*C", Fast Software Encryption. 7th International Workshop, FSE 2000. Proceedings (LNCSVol. 1978). Springer-Verlag. 2001, pp. 1-18. Berlin, Germany.

Briceno, Goldberg, Wagner, "*GSM Cloning*", http://www. isaac.cs.berkeley.edu/isaac/gsm-faq.html, 1998.

Ekdahl, P. Johansson, T. *"Another attack on A5*/1", IEEE International Symposium on Information Theory-Proceedings 2001. pp. 160 (IEEE cat. n 01CH37252)

Brookson, GSM (and PCN) *Security and Encryption*, 1994, http://www.brookson.com/gsm/gsmdoc.htm.

M. Walker and T. Wright, *Security*. In F. Hillebrand, editor, *GSM and UMTS: The Creation of Global Mobile Communication*, pp. 385-406, John Wiley & Sons, New York, 2002.

R. J. "Bud" Bates, *GPRS*, McGraw Hill TELECOM, 2002.

Jörg Eberspächer and Hans-Jörg Vögel. *GSM switching, services and Protocols*. John Wiley and Sons, 1999.

Garg, Vijay K. *Principles and applications of GSM*. Upper Saddle River (NJ) Prentice Hall PTR, 1999.

ETS 300 534. *Digital Cellular Telecommunication System (Phase 2); Security Related Network Functions*. ETSI, August 1997.

ETSI TS 100 929. *Digital Cellular Telecommunication System (Phase 2); Security related network functions*. ETSI, November 1999.

Lauri Pesonen, *GSM Interception*, http://www.dia. unisa.it/ads.dir/corsosecurity/www/CORSO-9900/a5/Netsec/netsec.html#chap1, Nov1999.

Bruce Schneier, *Applied Cryptography: Protocols, Algorithms, and Source Code in C*, 2nd Edition, Wiley, 1995.

N. Doraswamy and D. Harkins, "*IPSec: The New Security Standard for the Internet, Intranets, and VPN*", PH PTR, 1999.

O. Shaham, S. Aviran, E. Simony, Y. Shapira, "*Rate Control for Advanced Wireless Networks*", http://www.wisdom.weizmann.ac.il/~odedsh/

M. Meyer, *TCP Performance over GPRS*, In Proc. of IEEE WCNC, 1999, http://www.cs.helsinki.fi/u/gurtov /reiner/wcnc99.pdf

RFC 2406

R. Ludwig and B. Rathonyi, *Link Layer Enhancement for TCP/IP over GSM,* Proceedings of the IEEE INFOCOM '99, April, pp. 415-422.

PRACTICAL AUDITABILITY IN TRUSTED MESSAGING SYSTEMS

Miguel Reis and Artur Romão
Novis Telecom, SA
Estrada da Outurela, 118-A, Carnaxide, Portugal
{*mreis,ar@isp.novis.pt*}

A. Eduardo Dias
Universidade de Évora
Rua Romão Ramalho, 59, Évora, Portugal
aed@di.uevora.pt

Keywords: Auditability, Trusted Repository, Secure Messaging, Secure Electronic Commerce.

Abstract: The success of a dispute resolution over an electronic transaction depends on the possibility of trustworthily recreating it. It is crucial to maintain a trusted, thus fully auditable, repository to which a judge could request a transaction recreation. This article presents a practical scheme providing strong guarantees about the auditability of a trusted repository. We use the messaging paradigm to present the mechanism, but it can be applied to any other scenario that needs to maintain fully auditable long term information.

1 INTRODUCTION

Common messaging systems, in particular electronic mail, do not possess enough security guarantees to satisfy most of the assumptions required on security demanding areas, such as military or business to business electronic commerce. Even secure electronic mail is not enough, since it only satisfies some requirements, like authentication, integrity, confidentiality and non-repudiation of origin.

Stronger security requirements, like non-repudiation of submission and non-repudiation of receipt (Kremer et al., 2002; Zhou, 2001), together with trusted auditability (Haber and Stornetta, 1997; Peha, 1999) from the message transportation and delivery systems, are not guaranteed at all. Furthermore, it is fundamental to assure the effective message delivery, or some warning about the impossibility of delivery, as well as reliable and secure (e.g., confidential) message archiving, needed for legal effects and long term availability.

This article presents an approach to maintain long term auditability of this type of electronic messaging systems, and it is organized as follows. In section 2 we present a series of required assumptions. In sections 3 and 4 we propose a new scheme and in section 5 we analyze its security. In section 6 we analyze the efficiency of the proposed scheme. In section 7 we present implementation guidelines using widely available tools. In section 8 we conclude the article.

2 SECURITY REQUIREMENTS

Messaging systems auditability is based on the possibility of recreating a transaction or a transaction set. This requirement demands the trusted storage of the set of messages belonging to a transaction. Every message, as well as additional attributes, is mapped to a specific record. A record is the basic unit of a trusted repository. We can define trusted storage as a series of assumptions made over a record:

- **Content integrity**: It is impossible to corrupt the content of a record without detection.
- **Temporal ordering**: Every record must be in chronological order, and this ordering cannot be corrupted without detection.
- **Record elimination**: It is impossible to delete a record without detection.
- **Record insertion**: It is impossible to add a non-authorized record without detection. We define authorized entity as someone possessing or having access to a secret needed to create records.

From this point on we will use the word validity" to refer to a situation where all the assumptions are achieved.

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 135–140.

3 A NEW AUDITABILITY SCHEME

In this section we describe a new scheme providing strong guarantees of meeting all the assumptions previously identified. Let us assume that all the records are kept inside a trusted repository. Let us also assume that message transportation from the messaging system to the trusted repository is done without corruption.

3.1 Notation

We use the following notation to represent data elements and functions in this article:

M: message belonging to a specific transaction

E: record element

E_1, E_2: concatenation of two elements E_1 and E_2

R = $\{E_1,...,E_n\}$: record containing the concatenation of elements E_1 to E_n

f_m, f_e, f_h: flags indicating the purpose of a record

L : label linking a message with a specific transaction (transaction identifier)

$H_k(E)$: keyed message digest applied to element E using key k

$H(E)$: message digest applied to element E

$s_k(E)$: signature applied to element E using private key k

V_A e S_A: verification and signature key of principal A

$E_{\langle n \rangle}$: element placed in position n in an ordered list

$T(E)$: timestamp (Adams et al., 2001) applied to element E

3.2 The protocol

Whenever a message M is sent to the trusted repository a new record R_m is created in the following way:

$$R_m = \{f_m, L, M, Mac\}$$

$$Mac = H_k(f_m, L, M)$$

The Mac element is generated using elements present in the record. If any of these elements changes, the Mac element must change too, in order to keep the record integrity. This way we ensure that only who owns the secret k is able to change or add records to the repository. With this mechanism we satisfy the integrity assumption.

We now introduce the concept of an *epoch*. An epoch is defined as a set of R_m records, ordered in time, and completed with an R_e record. This type of record is defined in the following way:

$$R_e = \{f_e, k, V_A, Sig_e, T(Sig_e)\}$$

$$Sig_e = s_{S_A}(f_e, k, H(Mac_{<1>}, ..., Mac_{<n>}))$$

Sig_e is a digital signature made over a sequence of elements belonging to the R_e record together with a message digest element. The message digest is built from an ordered sequence of elements belonging to all R_m records which form this epoch.

As explained above using R_m records, R_e records can only be changed or added to the trusted repository by who owns a secret, which is, in this particular case, the signature key S_A. By using a message digest built over elements orderly collected from all the records R_m included in this epoch, the signature element Sig_e gives us the guarantee of content integrity, temporal ordering, non-elimination and non-authorized insertion of records without detection.

The message digest function referred above is created using Mac elements. This way we not only guarantee the integrity of these particular elements within each R_m record, but also the integrity of the set of R_m records belonging to this epoch.

So far we only guarantee the assumptions defined in section 2 within each particular epoch (as long as it is closed). But we must also guarantee that epochs are ordered in time, as well as the impossibility to completely remove one or more epochs without detection, thus certifying that the assumptions defined in section 2 are verified in all the trusted repository. To achieve this goal we need to modify the definition of the Sig_e element in the following way:

$$Sig_{e_{<n>}} = s_{S_A}(f_e, k, H(Mac_{<1>}, ..., Mac_{<n>}), H(Sig_{e_{<n-1>}}))$$

This way all of the R_e records are directly connected and ordered in time. Every R_e record has among its elements a reference to the immediately

previous R_e record (as depicted in figure 1). For someone to be able to compromise one or more epochs without detection from the validation system, all of the epochs generated after those ones would also have to be changed. Assuming the signature key used in the last R_e record of the trusted repository is never compromised, we can state that with this scheme all of the assumptions defined in section 2 are fully satisfied.

R_e records act as checkpoints of validation throughout the repository. The concept of creating checkpoints and link them together follows the concept of a Merkle tree (Merkle, 1980).

Figure 1: Example with a non-fixed number of R_m records in each epoch.

3.3 Record Integrity Verification

To validate the content of an R_m record it is necessary to go through the following steps:

1. Check which epoch the R_m record belongs to, thus identifying that epoch's R_e record.
2. Obtain $H(Mac_{<1>}, ..., Mac_{<n>})$. Mac elements are obtained from every R_m record belonging to the current epoch.
3. Obtain the previous epoch R_e record, using it to obtain $H(Sig_{e_{<n-1>}})$.
4. Check if the content from record R_e is not corrupted, using verification key V_A, the elements gathered in the previous steps and elements f_e and k to validate the digital signature present in element Sig_e.
5. Validate the content of record R_m by checking if the content from Mac element matches $H_k(f_m, L, M)$. We need to use the secret element k present in this epoch's R_e record.
6. Repeat steps 2 to 4 to all of the epochs created after this one, thus validating the chain of R_e records until the last one currently present in the trusted repository is correctly verified, or an error occurs.

4 HIERARCHICAL PARTITIONING OF RECORDS

The procedure explained above becomes impractical as the number of epochs increases. This is due to the fact that the number of record verifications that are necessary is directly proportional to the number of records in the trusted repository.

4.1 Definitions

To solve this problem we introduce a new hierarchical partitioning scheme. Instead of directly connect all the R_e records in the trusted repository, we now only directly connect R_e subsets. To explain this scheme we present a new notation:

$R_{[y]}$: record belonging to hierarchical level y

$R_{<l>}$: last record of a subset

The last record of an epoch subset, which is always an R_e record, is no longer directly connected to the first R_e record of the next epoch. Instead, it is now only directly connected to a hierarchically superior record. This new type of record will be defined as R_h. R_h records are also grouped in subsets, and directly connected one to another, just like explained to R_e records. In a generic way, every time a record subset is terminated with an $R_{h_{[y]<l>}}$ record, a new hierarchically superior record level beginning with an $R_{h_{[y+1]<1>}}$ record is created. The first record belonging to hierarchical level $y+1$ is always directly connected to the last record belonging to hierarchical level y (as depicted in figure 2). We now formally introduce this new type of record:

$R_{h_{[y]}} = \{f_h, V_A, Sig_{h_{[y]}}, T(Sig_{h_{[y]}})\}$, where

$Sig_{h_{[1]<n>}} = s_{S_A}(f_h, H(Sig_{e_{<l>}}))$

$Sig_{h_{[y]<1>}} = s_{S_A}(f_h, H(Sig_{h_{[y-1]<l>}}))$ if $y \neq 1$

$Sig_{h_{[y]<n>}} = s_{S_A}(f_h, H(Sig_{h_{[y]<n-1>}}))$ on all other situations

4.2 Integrity Verification Procedure

With this approach, to validate the content of an R_m record we begin by going through all the steps defined in section 3.3 with some minor changes. We re-define the last step and extend the procedure:

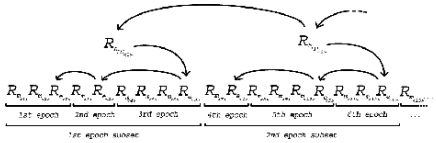


Figure 2: Hierarchical record scheme example.

6. Repeat steps 2 to 4 to all of the epochs created after this one and belonging to the current subset, thus validating a chain of R_e records.
7. Find which record $R_{h_{[1]}}$ is directly connected to the last record of the current subset, $R_{e_{<l>}}$, and check its content integrity by validating the signature in element $Sig_{h_{[1]}}$. Check also the integrity of all the other $R_{h_{[1]}}$ records belonging to the same subset.
8. Find which record $R_{h_{[y+1]}}$ is directly connected to the last record of the current subset, $R_{h_{[y]<l>}}$, and check its content integrity by validating the signature in element $Sig_{h_{[y+1]}}$. Check also the integrity of all the other $R_{h_{[y+1]}}$ records belonging to the same subset.
9. Repeat the previous step until the hierarchically highest level as been successfully verified.

With the procedure explained above, verifying the integrity of some R_m record is no longer directly proportional to the number of existing records, as explained in section 6.

5 SECURITY ANALYSIS

The integrity of some record R_m is based on the security of the underlying message digest algorithm, as well as on the privacy of the secret k used to calculate Mac elements. Due to this fact, we should use a well known message digest algorithm, whose inviolability is perfectly demonstrated. We should also choose a secret in line with the computational capabilities available during the underlying epoch, minimizing the risk of well succeeded attacks over Mac elements. To prevent an attacker from manipulating R_m records, it is vital to keep the privacy of secret k assured as long as the current epoch is not completed with the generation of an R_e record.

The secret k used to build Mac elements is present in the respective R_e record. This is not a security weakness, since the epoch is closed and its security now lies on the integrity of the Sig element.

The content integrity of record types R_e and R_h is based on the security of the signature algorithm, as well as on the secrecy of the signature key. This leads us to conclude that epoch validity is dependent on the Sig element integrity. If the signature key becomes compromised the current epoch becomes also compromised, due to the possibility of re-signing it without detection.

In order for this compromise procedure to become fully undetectable it is also necessary to compromise the chain of directly connected records. This implies compromising all the R_e records belonging to the current subset and created later in time, as well as all the directly connected and hierarchically superior R_h record subsets (as depicted in figure 3). We can conclude this from the fact that every record belonging to the chain has among its elements a reference to the Sig element of some previously created record.

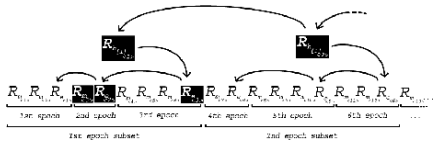


Figure 3: Undetected corruption of record $R_{m_{<3>}}$ implies compromising the chain of directly connected records.

This fact, together with the possibility of renewal of signature keys from one epoch to another makes undetected corruption extremely hard to achieve. Since signature key sizes can (and should) be adapted to the current computational power, the greatest security concern comes from the ability to maintain the secrecy of the signature key itself. We should eliminate the signature key after the end of an epoch, minimizing the risk of key disclosure. Since we are using asymmetric cryptography, this key is no longer necessary to verify a record's validity.

Nevertheless, it is possible that signature keys used a long time ago will become compromised (i.e., discovered), mostly due to technology breakthroughs. Even in this case, the record integrity may remain fully verifiable. All we need to do is to check if the T element remains valid. This element holds a timestamp of the Sig_e or Sig_h element, thus placing an upper boundary on the date the signature has been made. If this timestamp is later than the date the signature key has been or known to be compromised, the record is marked as valid. The T element is created by applying a digital signature to the target data. We assume the private key used to generate this signature is not compromised.

6 EFFICIENCY

Adapting secret k or signature key sizes to the current computational power is a basic operation, since both can be replaced after every epoch.

Creating R_e and R_h records takes much more time than creating R_m records, mostly due to the characteristics of asymmetric versus symmetric cryptography (Schneier, 1995). To minimize this constraint it is possible to increase the number of R_m records per epoch, keeping always in mind the fact that one must never allow the computational power available during the current epoch to be able to compromise secret k used to build Mac elements.

Secret k is revealed in the end of each epoch, thus making a validation over an R_m record a very easy and fast operation to conclude, even after long time periods. Verifying the validity of an R_m record implies the validation of $O(\log_x n)$ records, where

n = total number of records in the trusted repository

x = average number of records within each subset

7 IMPLEMENTATION GUIDELINES

In this section we present guidelines to an implementation of the scheme defined in the previous sections, through the application of technology widely used and proved to be secure and efficient nowadays.

The message digest algorithm used to create Mac elements must be widely deployed and proved to be secure, and at same time efficient, since it will be used very often. A keyed-hashing algorithm like $HMAC$ (Krawczyk et al., 1997) satisfies all of the above requirements.

In a similar way, it is important to use a widely deployed one-way hash function like $SHA-1$ (NIST, 1994) to generate elements which will be needed to create Sig_e or Sig_h elements.

Sig_e and Sig_h elements will be built using asymmetric cryptography. The private key is used to generate those elements and the public key, which will be bound to a $X.509$ (ITU-T, 2000) digital certificate, is used to verify the integrity of the elements, interacting with a certification authority (CA). Despite the fact that the trusted repository must fulfill certain security requirements, critical operations like certificate life cycle management are much more suitable to be done by a CA.

It is crucial to the preservation of the protocol security to establish a security scheme for the certificate requests (RSA, 2000) to be done. There are several good approaches: one is to use a mutually-authenticated TLS (Dierks and Allen, 1999) connection. Other may be by carrying some shared secret, which should be evolving from one request to the next, among the certificate request extensions.

Another point where we may increase security is by settling an agreement with the CA in which we can define a set of $X.509$ extensions to be included in all the certificates issued to the service. By issuing specific extension values for each digital certificate we may, for instance, lower the risk of certificate replacement frauds.

Validating an R_m record must also always require validating the digital certificate state by using an CRL (ITU-T, 2000) or an $OCSP$ (Myers et al., 1999) service. Whenever a certificate is found to be invalid (revoked, expired, etc.) it is necessary to validate the timestamp present in the T element, as explained previously, to decide if the record remains valid.

One final note concerning secret and private key protection. As pointed out before, it is extremely important to keep the items private. To achieve this goal we should use cryptographic hardware which allows us to generate secret and private keys inside an hardware token. The keys also have the possibility to never leave the token, thereby creating a very secure environment.

8 CONCLUSION

In this article we proposed a new scheme providing strong guarantees about the auditability of a trusted repository. The trusted repository maintains three types of records:

- R_m records keep the actual messages belonging to a specific and well defined transaction.
- R_e records aggregate sets of R_m records together, establishing epochs. This type of records are also aggregated in sets. In every set an R_e record keeps a reference to the R_e record created immediately before.

- R_h records are aggregated in directly connected sets and bound to a hierarchical level. Generically, every first R_h record of a subset belonging to hierarchical level y holds a reference to the last R_h (or R_e if y represents the first hierarchical level) record of a subset belonging to hierarchical level $y - 1$.

We state that maintaining the secret k used to build R_m records private as long as the current epoch is not terminated, and adapting the size of this secret to the computational power available during the present time makes undetected corruption of R_m extremely records hard to achieve. Besides that, if we maintain strong guarantees that the private keys used in creating the last record of every hierarchical level are not compromised, undetected corruption of the complete chain of R_e and R_h records alse becomes extremely hard to achieve.

We have provided guidelines that prove informally that it is possible to make a practical implementation of this scheme through the application of technologies available in the present time. Although we use the trusted messaging system concept as a way of presenting the scheme, it can be applied to any system that needs to maintain fully auditable long term information.

REFERENCES

Adams, C., Cain, P., Pinkas, D., and Zuccherato, R. (2001). Time-stamp protocol (tsp). RFC 3161, Internet Engineering Task Force.

Dierks, T. and Allen, C. (1999). The tls protocol version 1.0. RFC 2246, Internet Engineering Task Force.

Haber, S. and Stornetta, W. S. (1997). Secure names for bit-strings. In *ACM Conference on Computer and Communications Security*, pages 28–35.

ITU-T (2000). Itu-t recommendation x.509. Technical report, ITU-T.

Krawczyk, H., Bellare, M., and Canetti, R. (1997). Hmac: Keyed-hashing for message authentication. RFC 2104, Internet Engineering Task Force.

Kremer, S., Markowitch, O., and Zhou, J. (2002). An intensive survey of fair non-repudiation protocols. *Computer Communications*, 25(17):1606–1621.

Merkle, R. C. (1980). Protocols for public key cryptosystems. In *Proceedings of the IEEE Symposium on Research in Security and Privacy*, pages 122–134. IEEE Computer Society Press.

Myers, M., Ankney, R., Malpani, A., Galperin, S., and Adams, C. (1999). X.509 internet public key infrastructure online certificate status protocol - ocsp. RFC 2560, Internet Engineering Task Force.

NIST (1994). *NIST Federal Information Processing Standard Publication 180-1: Secure Hash Standard.*

Peha, J. M. (1999). Electronic commerce with verifiable audit trails. In *Proceedings of ISOC*.

RSA (2000). Pkcs #10 v1.7: Certification request syntax standard. Technical report, RSA Laboratories.

Schneier, B. (1995). *Applied cryptography: protocols, algorithms, and source code in C*. John Wiley and Sons, Inc., second edition.

Zhou, J. (2001). *Non-Repudiation in Electronic Commerce*. Artech House, first edition.

TOWARDS AN ADAPTIVE PACKET MARKING SCHEME FOR IP TRACEBACK

Ping Yan and Moon Chuen Lee
Department of Computer Science and Engineering, The Chinese University of Hong Kong, CUHK, Hong Kong
Email: pyan@cse.cuhk.edu.hk, mclee@cse.cuhk.edu.hk

Keywords: DoS attacks, IP traceback, probabilistic packet marking, inter-domain marking, source router marking, attack graph reconstruction.

Abstract: Denial of Service attacks have become one of the most serious threats to the Internet community. An effective means to defend against such attacks is to locate the attack source(s) and to isolate it from the rest of the network. This paper proposes an adaptive packet marking scheme for IP traceback, which supports two types of marking, namely source router *id* marking and domain *id* marking. For each packet traversing, we let the border routers perform probabilistic router *id* marking if this packet enters the network for the first time, or perform probabilistic domain *id* marking if the packet is forwarded from another domain. After collecting sufficient packets, the victim reconstructs the attack graph, by which we keep track of the intermediate domains traversed by attack packets instead of individual routers within a domain; however, the source routers serving as ingress points of attack traffic are identified at the same time. Simulation results show that the proposed marking scheme outperforms other IP traceback methods as it requires fewer packets for attack paths reconstruction, and can handle large number of attack sources effectively; and the false positives produced are significantly low. Further, it does not generate additional traffic.

1 INTRODUCTION

The intent of *Denial of Service* (DoS) attacks is to prevent or impair the legitimate use of computer or network resources. Internet connected systems face a consistent and real threat from DoS attacks, because the Internet fundamentally composed of limited and consumable resources like bandwidth, processing power, and storage capacities is rather vulnerable to some level of service disruption (Kevin J. Houle et al., 2001). In case of *Distributed Denial of Service* (DDoS), an attacker first compromises a bunch of hosts weakly secured or possessing vulnerable network service programs, and he then uses these compromised computers to launch coordinated attacks on victim machines.

The primary difficulty of dealing with (D)DoS attack is *IP Spoofing*, which is almost always present in such attacks. In order to prolong the effectiveness of the attack, the attackers spoof the source IP addresses in their attacking packets to avoid being traced. Therefore, in a *traceback* problem, our task is to find out the actual source(s) of the attack, where we define the *source* as the router directly connected to the system from which the flow of packets, constituting the attack, was initiate (Hal Burch, 1999) (Steven H. Bass, 2001). Upon identifying the attack source(s), the victim or the network operators can conduct efficient defenses against DoS or DDoS attacks, either by blocking the traffic from the identified sources or filtering out the malicious packets on their way to the victim.

1.1 Problem Model and Performance Metrics

We would model the attack with a number of coordinated attackers attacking a single victim as an undirected graph with each node representing a domain. Domain is a logical subnet[1] on the Internet; a campus or internal corporate network is an example of a domain. Data exchange between campus and corporate domains is facilitated by one or more ISP domains, which offer, as a service, transmission and switching facilities for data exchange between their

[1]Subnet is a portion of a network, which may be a physically independent network segment, which shares a network address with other portions of the network and is distinguished by a subnet number (Tanenbaum, 2002).

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 141–148.

customers. The IP packets thus flow though different network domains, from regional ISP network to international ISP network and finally get to the destination. In general, to model the attack, a network domain can be thought of as a cloud, which connects to other domains at the peering points, with clients attaching on border routers. In our solution, the reconstructed *attack graph* would incorporate attack paths and the source router(s) identified, with each node on the paths can be viewed as a domain.

In the literature, the performance of IP traceback approaches is commonly measured by several parameters. *Minimum number of packets* is the number of packets required for attack graph reconstruction; it is desired to be minimized to achieve a fast response to an attack and diminish the damage (Savage et al., 2000) (Kuznetsov et al., 2002). A *false positive* is a router that is actually not on an attack path but is reconstructed to an attack graph by a traceback mechanism, and a *false negative* is a router that is missed in the reconstructed attack graph (Savage et al., 2000) (Kuznetsov et al., 2002). Furthermore, an efficient traceback approach should feature a relatively low *computation complexity* and incremental deployment into the current Internet structure, at low cost (Kuznetsov et al., 2002).

Our proposed method will not incur network traffic overload or storage overhead on the participating routers, though certain memory is required at the victim site. Therefore, we would assess the proposed method mainly based on the above parameters and the simulation results are demonstrated in section 4.

1.2 Overview Of the Proposed Method and Contributions

In this paper, we present a practical IP traceback approach. It addresses the issues concerned by both the victims and network operators such as per-packet marking space limitation, network overload and computation overhead.

In our proposed marking algorithms, we employ 25 bits space in the IPv4 packet header as marking fields. Probabilistically, each participating router adaptively inscribes onto a traversing packet with its local partial path information. There are two types of markings: router identification (*rid*) marking and domain identification (*did*) marking. The *rid* marking is executed if the packet enters the network for the first time; in contrast, *did* marking is performed when the packet traverses along the following domains towards the victim.

The victim under a DDoS attack reconstructs the attack graph in two phases. First, it identifies all the intermediate domains taking part in forwarding

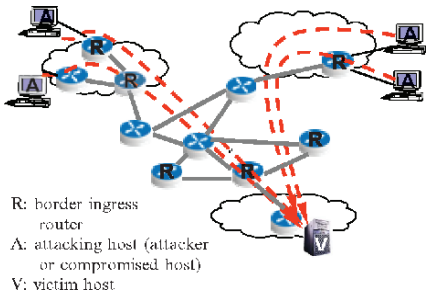


Figure 1: Adaptive traceback mechanism equipped network.

the attack packets, and recovers the inter-domain attack paths by inspecting the domain *ids* marked in the received packets. Second, from the router *id* markings in the received packets, the victim can identify the source routers. In general, the inter-domain attack paths reconstruction leads to the identification of the *source domains* (where the source router may reside), and then the identified source domains are associated with the router *id* markings to uncover the source routers as we wish.

This work presents a novel design of probabilistic packet marking scheme at the granularity of domain, while at the same time, the attack sources can be identified. And this inter-domain IP traceback design has been proved to possess the following advantages. First, we use much less number of packets to identify the attack source(s). In particular, it requires only two uniquely marked attack packets to identify a source router. Second, this approach generates quite low false rates which will be demonstrated by the simulation results. Third, our traceback mechanism ensures incremental deployment and requires fewer routers to participate. Actually, only the border routers[2] need to take part in this traceback mechanism. Figure 1 shows the example network with the implementation of our traceback mechanism.

Furthermore, keeping track of the domains traversed instead of intermediate routers on the attack paths and using *ids* instead of full 32-bit IP addresses have some underlying advantages. First, a domain is an administrative unit on the Internet, which has the capability to conduct defenses against the attacks when it is identified to be involved in an attack and notified by the victim. Second, it still makes sense to identify the source domains even the individual source routers are not identified correctly. We observe that systems of some domains are more

[2]A router that sits on the border of a network connecting it to an end-host or a router in another network.

likely to be compromised (to launch a DDoS attack, attackers usually compromise a bunch of vulnerable systems as attack agents) due to the domains' poor security features such as weak intrusion defense mechanisms and flawed security policies.

1.3 Organization Of the Paper

In the second section, we survey the previous work on traceback problem. In section 3, we present our proposed traceback scheme in depth and articulate the operation of our traceback mechanism. At last, we give the simulation results in section 4. Section 5 concludes the paper.

2 RELATED WORK

2.1 Tracing Hop-by-hop

J. Ioannidis and S. M. Bellovin (Ioannidis et al., 2002) proposed a *Pushback mechanism*. In this approach, a congested router nearest to the victim uses statistics and pattern analysis to determine from which most adjacent upstream routers the unexpected traffic volume are coming, and then send signals to notify the traffic contributors to rate-limit the suspect traffic. The approach is then repeated at the upstream routers in a chain to identify and rate-limit the traffic contributors. This scheme therefore requires immediate action during the attack, and requires considerable coordination between network operators. The main drawback with this method is that, in large-scale DDoS attacks, they have limited capabilities to separate the legitimate packets from attack packets in a pattern-based way.

2.2 ICMP Traceback Messaging (iTrace)

S. M. Bellovin (Bellovin, 2000) proposed an alternative approach, *ICMP traceback messaging* (or simply iTrace). With some probability q (typically, q = 1/20000 is proposed), each router sends an additional ICMP message packet to the destination for each packet it received. The message contains information of the local router traversed and its adjacent hop. With sufficient ICMP traceback messages from routers along the path, the attack source(s) and paths can be determined at the victim site. The main drawback of this approach is that it causes additional network traffic even when no attack is present. Consequently, q should be small enough to imply a relatively low network traffic overload. However, using a small q, this approach is inefficient in terms of the number of ICMP traceback packets required. For example, if the maximum path length is 20 and there are about 1000 nodes on the reconstructed attack graph, the expected number of attack packets required to arrive at the victim to reconstruct the attack graph is 7.5 million (Goodrich, 2002).

2.3 Logging & Querying

In a logging solution, we let the routers log the packets they process, and a victim then actively queries the routers to see whether they sent suspect attack packets. In general, this approach is infeasible because of the huge storage requirement at the routers (Ioannidis et al., 2002). However, *Source Path Isolation Engine* (SPIE) has the capability of identifying the source of a particular IP packet given a copy of the packet to be traced, its destination, and an approximate time or receipt (Snoeren et al., 2002). Most notably, with the use of an innovative logging technique, collecting only the hashes of the packets, this approach reduces the memory requirement down to 0.5% of link bandwidth per unit time (Snoeren et al., 2002). However, though the storage requirement has been significantly reduced, the overhead is still considerable.

2.4 Probabilistic Packet Marking

To avoid the network overloading, some researchers propose to embed traceback information in the IP packets, which is commonly referred to as *probabilistic packet marking* (or simply PPM) method. Savage *et al*. (2000) proposed to let each router mark each packet it forwards with a piece of partial path information at a set probability p (e.g., p = 1/20). A message "edge" recording the identities of a router and its previous hop would be inscribed onto certain bits employed as marking fields in the IP header. However, the edge message has to be made to fit in the limited reserved bits; so they break it into fragments sent by separate packets. To reconstruct the attack paths, every possible fragments combination is tried to form a valid edge, and then the edge is used to recover the sequence of intermediate routers hop by hop at the victim site. Unfortunately, for even small-scale distributed DoS attacks, this method is not practical due to the tremendous combinatorial trials and the high false rates. Even worse, it would introduce many false positives because the previous mis-reconstructed messages lead to more false combinatorial trials, which can be described as "explosion effect".

In Advanced and Authenticated Marking Schemes, Song and Perrig (2001) proposed the use of hash chains for authenticating routers to improve the performance of probabilistic packet marking. They do not fragment router messages. Instead, they assume the victim knows the map of its upstream routers, so the full IP address is encoded into 11 bits hash values by two sets of universal random hash functions in the packet marking. To reconstruct the attack graph, the victim uses the upstream router map as a road-map and performs a breadth-first search from the victim to identify the corresponding router which was hashed and written into the marking fields.

3 ADAPTIVE PACKET MARKING SCHEME

Our adaptive packet marking scheme is based on the probabilistic packet marking technique, but a novel IP packet marking scheme is proposed, which is motivated by the below issues.

3.1 Design Motivation

The IP traceback approaches, such as iTrace or the proposed probabilistic packet marking schemes, rely on observing a high volume of spoofed traffic comprised of thousands or millions of packets, so the attacker can undermine the traceback by spreading the attack traffic across many attacking hosts (also referred to as agents, slaves, or reflectors in a reflector DDoS attack (Chang, 2002)), greatly increasing the amount of time required by the traceback scheme to gather sufficient packets to analyze. Therefore, an effective traceback scheme should use as few packets as possible to reveal an attack path. Using a relatively short *id* instead of a full IP address, we do not need to spread a mark across multiple packets, and we thus feature a relatively small number of packets to fulfill the traceback.

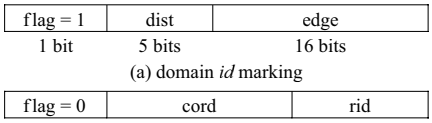
In addition, some people are challenging the necessity of the full-path traceback solution (Belenky et al., 2003); identifying all the intermediate routers that the attack packets traversed, may be unattractive to the victims and ineffective for DoS (DDoS) countermeasures. First, the full-path traceback is as good as the address of an ingress point in terms of identifying the attacker. Second, each packet in a datagram network is individually routed so packets may take different routes even if their source and destination are identical. Third, the addressing within ISPs' networks is not necessarily understandable to the public since ISP may use private addressing plans within their own networks (Belenky et al., 2003). Therefore, we propose a domain based IP packet marking scheme to identify the intermediate domains instead of the individual routers, except the one serving as the attack source. In the following paragraphs, we will describe the proposed scheme in depth and state how this method addresses the problems with the existing solutions.

3.2 Using ID for Marking

The proposed marking scheme overloads 25 bits space in IPv4 header; the 25 bits space consists of the 16-bit Fragment Identification field, 1-bit fragmentation flag and 8-bit Type of Service (ToS) field. Employing the 25 bits in the IP header for marking was first advocated by Dean et al. (2001). The ToS field is currently not set except for extreme unusual cases. The Fragment ID field is a 16-bit field used by IP to permit reconstruction of fragments; this field is commonly used as a marking field and the backward compatibility is fully discussed in Savage's paper (Savage et al., 2000). The fragmentation flag is an unused bit that current Internet standards require to be zero. We also see there are some proposals on marking in the IPv6 header; however, it is not to be discussed in this work.

As every host or router on the Internet is identified using a 32-bit IP address (Tanenbaum, 2002), it is a challenging issue to overload the 25-bit marking space in the IP header with a 32-bit IP address. In our proposal, since we only aim to identify the intermediate domains taking part in the attack and the source routers, there is no need to use full IP addresses, as long as we can uniquely identify each domain with a different identification. If we assign a 16-bit domain *id* to each domain, we can uniquely identify up to 2^{16} (65536) domains. If we assume there are at most 2^{10}(1024) border routers within a domain, a 10-bit value is sufficient to be assigned as a router *id* to identify the source routers within a source domain. However, to defend against the attack, the victim may demand to block the malicious traffic at the source routers, so the victim needs to retrieve the IP addresses from the *ids*. This could be implemented as an ID-to-IP mapping table published on websites, or it could be maintained at the victims individually.

Two types of markings, either the router *id* marking or the domain *id* marking, are to be performed by a router adaptively by checking whether the router concerned is the ingress point of the to-be-marked packet or not. At first, however, the border routers, with implementation of our marking scheme should be capable of determining which type of marking to perform. Physically, these routers



(a) domain *id* marking

(b) router *id* marking

Figure 2: Marking fields.

are connected to end-hosts or other routers through different interfaces; a router therefore checks through which interface it receives a packet to see whether a packet is forwarded by another router from outside the domain concerned or sent by an end-host at the customer side. The domain *id* marking would be performed if the router receives a packet routed from outside the current domain, and the router *id* marking would be performed when the packet comes from an end-host. Figure 2 shows the marking fields for domain *id* marking and router *id* marking respectively.

3.3 Domain id Marking

The domain *id* marking algorithm allows the victim to infer the inter-domain attack paths by inspecting the domain *ids* in the received packets. As shown by Figure 2 (a), "edge" field stores one encoded edge on an attack path, and the 5-bit distance field represents the number of hops traversed since the edge it contains is sampled. A flag is used to indicate whether this is a domain *id* marking or a router *id* marking. Basically, the domain *ids* of two neighboring domains are encoded by *exclusive-or* (XOR) to make up the edge, and it can be decoded back during reconstruction in virtue of XOR's property that $\alpha \oplus \beta \oplus \alpha = \beta$. This XOR encoding technique is used to reduce per-packet storage requirement.

Figure 3 shows the domain *id* marking scheme. Marking probability *p* determines whether to mark a packet or not. To mark the packet, router *R* sets the distance to be zero and writes the domain *id* into the edge field. Otherwise, if the distance is zero, router *R* overwrites the edge field with XOR of edge value present in the to-be-marked packet and its own *id*. The distance field is used as hop counts, and is always incremented, which is critical to minimize the spoofing of the markings by an attacker, so that a single attacker is unable to forge an edge between itself and the victim (Savage et al., 2000). Repeatedly, this procedure takes place for the following domains as the packets traverse along the path. We also remark that incremental deployment is ensured, because we can identify a domain even if only one border router within that domain sees the attack packet and marks it by our marking scheme.

Algorithm 1 Domain id marking by border router R

for each packet *pkt* from an upstream domain **do**
  generate a random number x within [0..1]
  if $x < p$ **then**
    pkt.edge = *did*
    pkt.dist = 0
    $pkt.flag = 1$
  else
    if *pkt.dist* is 0 **then**
      pkt.edge = *pkt.edge* ⊕ *did*
    increment *pkt.dist*

Figure 3: Domain id marking algorithm.

3.4 Router id Marking

Figure 4 outlines the algorithm for router *id* marking by router *R*. The router *id* marking algorithm is used to identify the source routers that serve as ingress points of attack packets. A router performs router *id* marking with certain marking probability if it receives a packet from the customer side. Recall that we refer to the domain where a source router resides as source domain. To complete the inter-domain attack path, the source domain *id* should also be conveyed to the victim; so we make it equally likely to mark a packet with the router *id* or the domain *id* at the source router. In practice, we use a larger marking probability *q* in router *id* marking procedure, which is double of the probability *p* that we use in domain *id* marking. It's like flipping a coin to decide to mark with domain *id* or router *id*. To mark a packet with a domain *id*, we set flag to be one and write the domain *id* into the edge field; and to mark the packet with a router *id*, we set flag to be zero and write the 10-bit router *id* into the rid field.

We also note that a 10-bit router *id* can be used to identify a router uniquely only within a domain; so we need to combine the router *id* and the corresponding source domain *id* to uniquely identify the source routers universally. We therefore write a 14-bit checksum *cord* side by side with the router *id* into the marking fields to associate a router *id* with the corresponding source domain. The checksum can be hashed from the 16-bit domain *id*, and it is sufficient for distinguishing 2^{14}(16384) source domains possibly involved in an attack. Therefore, we can place the identified source routers in their corresponding source domains according to the checksums to complete the attack graph reconstruction.

Algorithm 2 Router id marking by router R

```
for each packet pkt passing through R do
   generate a random number x within [0..1]
   generate a random number r within [0..1]
   if x < q then
      if r < 0.5 then
         pkt.edge = did
         pkt.distance = 0
         pkt.flag = 1
      else
         pkt.rid = rid
         pkt.cord = hash(did)
         pkt.flag = 0
```

Figure 4: Source router id marking algorithm.

3.5 Attack Graph Reconstruction

Figure 5 describes the reconstruction procedure; it's a two-phase procedure. The first one is inter-domain attack graph reconstruction using packets marked with domain *ids* and then by the end of first stage, the source domains are identified. In the second phase, the algorithm relies on the packets with router *id* markings to identify the source routers. The property of XOR that $\alpha \oplus \beta \oplus \alpha = \beta$ allows us to decode the domain *ids* hop-by-hop during inter-domain attack path reconstruction. The algorithm starts from recovering the domains one hop away from the victim. Let the victim's domain *id* be α and a domain closest to the victim have an *id* β. By $\alpha \oplus \beta \oplus \alpha = \beta$, the domain *id* β can be decoded, given the attack packet marked with the value: $\alpha \oplus \beta$, and distance equal to 1. Likewise, for all domain-*id*-marked packets at distance d, pkt_i, a number of candidate domain *ids* can be generated by XORing the edge it contains with the *ids* of previously reconstructed domains at distance d−1. We denote a candidate as D_{ij}, which is decoded from pkt_i and the known endpoint of the edge is node D_j. Then the victim checks the upstream domain topology, *M* as a road-map to verify candidate D_{ij} by checking if an edge does exist between the candidate and node D_j on *M*. Node D_j and the verified candidate D_{ij} would therefore make up an edge on the reconstructed attack path. Hop by hop, the inter-domain attack graph is thus reconstructed by the repeated process.

To locate the source domain associated with checksum C_1 (let the checksum marked in an attack packet be C_1), the victim performs a breadth-first search from the victim on the reconstructed inter-domain attack graph, level by level until it gets to the node with a checksum equal to C_1. Suppose we denote a source domain node as *SD* and denote the

Algorithm 3 Path reconstruction at victim V

```
let max_d be maximum attack path length
let G be reconstructed attack graph, initialized with the
vertex V
let M be the upstream inter-domain Internet map
Let SD.Rset be the set of source routers in domain SD

//Inter-domain attack graph reconstruction
for d = 1 to max_d do
  for each node D at distance d − 1 in G do
    for each packet pkt with distance d, flag 1 do
      candidate = pkt.edge ⊕ D.did
      if candidate is equal to D'. did and D' is one of
      the upstream nodes of D on M then
        insert a new edge (D , D') into G

//Source router identification
for each attack packet pkt with flag 0 do
  for each node SD in G do
    if hash(SD.did) is equal to pkt.cord then
       insert pkt.rid into SD.Rset

output all the attack paths in G
```

Figure 5: Attack path reconstruction algorithm.

set of source routers that may belong to domain *SD* as *SD.Rset*. The victim first sorts the router-*id*-marked packets by their cord field, finds out the packets with a matched checksum of *SD*, and then adds their router *ids* into *SD.Rset* (it is initially empty and it does not include any duplicates). The victim thus reconstructs the attack graph that incorporates the identified intermediate domains and the source routers.

4 PERFORMANCE EVALUATION

To test the performance of our proposed marking scheme, we conduct a number of experiments using an Internet map based on the traceroute dataset of the real Internet from CAIDA's Skitter Internet mapping project (CAIDA, 2004); the dataset contains 178,207 distinct traceroute paths widely distributed over the entire Internet. For the experiments, we use the first two bytes of an IP address as a domain *id* used in the domain *id* marking. On the other hand, the last two bytes of an IP address would be processed to be used as router *id* which would be used in the router *id* marking. The experiments are performed to assess the performance of our marking scheme characterized by a number of parameters, namely the minimum number of packets for the reconstruction of an attack path of a certain length, false positives, number of attack sources, and attack path reconstruction time. The results are as presented in Figures 6 to 8. Each data point in Figures 6

to 8 corresponds to an average value based on around 1000 experiment runs.

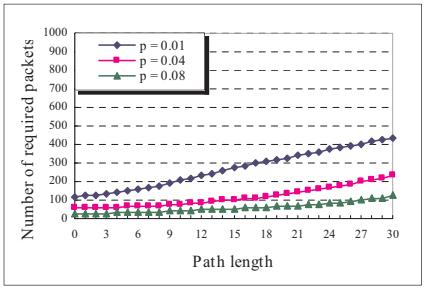


Figure 6: Number of packets required for attack paths reconstruction for different path lengths and different marking probabilities.

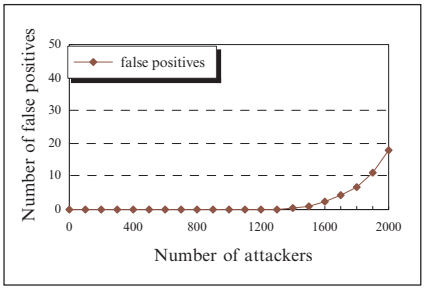


Figure 7: False positives generated for different number of attackers.

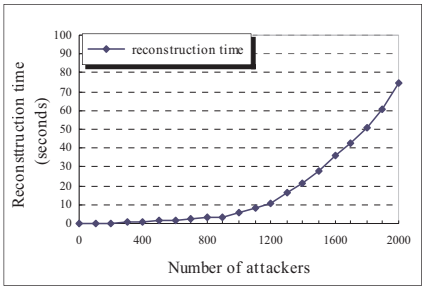


Figure 8: Reconstruction time for different number of attackers.

Figure 6 shows the minimum number of packets required for the reconstruction of attack paths of different lengths and different marking probabilities. Since a packet will normally traverse no more than 30 routers in the Internet to reach its destination, the attack path lengths considered in the experiments range from 0 to 30. In general, for each marking probability, the number of required packets for path reconstruction increases linearly with the path length. For the case with marking probability 1%, and path length 30, the required number of packets would be around 400; if the marking probability is 4%, roughly 200 packets would be required. When compared with other IP traceback methods, our proposed marking scheme requires fewer packets for reconstruction. For instance, for a marking probability of 4% and path length 30, scheme 2 with $m > 5$ and scheme 2 with $m > 6$ of the Advanced Marking Schemes (AMS) (Song et al., 2001), around 1000 and 4000 packets respectively are required for reconstruction, where m is the number of hash functions used to encode the router identification.

If we conservatively assume each domain contains two routers, a maximum domain based path length would be half of its equivalent router based path length. This implies that our marking scheme needs to handle attack paths with average path length equal to one half the path lengths of those handled by other marking schemes. Moreover, we need only one marked packet to identify a domain; whereas other marking schemes normally employ full *IP* address markings for full-path reconstruction, and several packets are usually required to identify each router on an attack path. For instance, eight marked packets are employed in both Savage's method (Savage et al., 2000) and Song and Perrig's method (Song et al., 2001) to encode each router's identity. So our marking scheme needs substantially fewer packets for attack graph reconstruction.

Figure 7 presents the number of false positives for different number of attackers in the range 100 to 2000. It shows that our marking scheme is free of any false positives even in presence of 2000 attackers. Though hash is used to encode the checksum which is used to locate source routers, 2^{14} (16384) different values are sufficient for the number of do mains possibly involved as source domains, that is , there is sufficient mapping space for the hashed values so that a collision-free hash function should be able to generate a near-zero result. For comparison, scheme 2 of Song and Perrig's method (Song *et al*., 2001) with m > 7 produced around 20 false positives in presence of 2000 highly distributed attackers. While our marking scheme has a computation complexity of around $O(dn^2)$, the method of Savage, *et* et al. (2000) and the method of Song and Perrig (Song et al., 2001) have a complexity of around $O(dn^8)$ and $O(dn^2)$ respectively, where d is the maximum path

length and n is the number of attacking hosts. Since our domain based marking scheme involves a smaller distance d, its complexity is relatively small.

Figure 8 presents the reconstruction times of our reconstruction algorithm for different number of attackers, measured on a 1500MHz Pentium IV PC platform. The results show that in general the attack graph reconstruction could be completed quite rapidly. Even for the case of 2000 attackers, it takes only about 50 seconds to reconstruct the attack graph, which is considered quite a fast response to a highly distributed large-scale DDoS attack.

5 CONCLUSION

This paper proposes an innovative marking scheme which supports two kinds of packet marking: inter-domain marking and source router marking. Based on the markings in the received packets, the victims can reveal the inter-domain attack paths and identify the source routers serving as the ingress points of attack traffic.

The advantages of the proposed IP traceback method include: (1) As the marking algorithms involve only the border routers, it ensures a practical implementation without a universal deployment on all the routers. (2) We keep track of domains traversed by attack packets other than all individual routers; as a result, attack paths reconstruction could be carried out more rapidly by our marking scheme. (3) Using the relatively short *id* instead of a full IP address, we do not need to split the markings for each domain into a number of fragments and the whole marking can be written into a single IP header. The number of packets required to identify each domain can thus be kept to a minimum.

Through the simulation experiments on the proposed marking scheme, we observe the following: (i) It requires a much smaller number of packets for attack paths reconstruction than other methods such as AMS (Song et al., 2001); (ii) It can handle multiple attack sources effectively in very large scale; (iii) The number of false positives generated even in the presence of 2000 attack sources is relatively small; (iv) It performs attack paths reconstruction quite rapidly and takes only around 50 seconds to reconstruct as many as 2000 attack paths on a Pentium IV PC platform. Thus it could be used to locate attack sources in real time, which is one of the critical steps in defending against DDoS attacks.

REFERENCES

Alezxx C. Snoeren, Craig Partridge, Luis A. Sanchez, Christine E. Jones, Fabrice Tchakountio, 2002. Single-Packet IP Traceback. *IEEE/ACM transactions on NETWORKING, Vol. 10, No. 6, December.*

Andrew S. Tanenbaum, Aug 9, 2002. *Computer Networks*, 4rd edition. Published by Prentice Hall PTR.

Andrey Belenky, Nirwan Ansari, 2003. IP Traceback with Deterministic Packet Marking, *IEEE COMMUNICATIONS LETTERS, Vol. 7, No. 4, APRIL.*

The Cooperative Association for Internet Data Analysis, 2004. Available: *http://www.caida.org/tools/measurement/skitter*

Dawn X. Song and Adrian. Perrig, 2001. Advanced and Authenticated Marking Schemes for IP Traceback. *Proc. of the IEEE Infocom conference, April.*

Hal Burch and Bill Cheswick, 1999. Tracing Anonymous Packets to Their Approximate Source. *Unpublished paper, December.*

J. Ioannidis and S. M. Bellovin. 2002. Implementing Pushback: Router-based Defense against DDoS Attacks. *Proc. in Network and Distributed System Security Symposium, the Internet Society.*

Kevin J. Houle, George M. Weaver, 2001. Trends in Denial of Service Attack Technology. *Technical report from CERT Coordination Center.* October.

Michael T. Goodrich, 2002. Efficient Packet Marking for Large-Scale IP Traceback. *CCS'02, November, Washington, DC, USA.*

Rocky K. C. Chang, 2002. Defending against Flooding-based Distributed Denial-of-service Attacks: a Tutorial, *IEEE Communications Magazine, October.*

S. M. Bellovin, 2000. ICMP Traceback Messages.

Internet Draft: *http://www.research.att.com/~smb/papers/draft-bellovin-itrace-00.txt* (June 20, 2004)

Stefan Savage, David Wetherall, Anna Karlin and Tom Anderson, 2000. Practical Network Support for IP Traceback. *Proc. of the ACM SIGCOMM conference, August.*

Steven H. Bass, 2001. Spoofed IP Address Distributed Denial of Service Attacks: Defense-in-Depth. Available: *http://www.sans.org/rr/papers/60/469.pdf* (July 30, 2004)

Vadim Kuznetsov, Andrei Simkin, Helena Sandström, 2002. An Evaluation of Different IP Traceback Approaches. Available: *http://www.sm.luth.se/csee/csn/publications/ip_traceback.pdf*

BASELINE TO HELP WITH NETWORK MANAGEMENT

Mario Lemes Proença Jr., Camiel Coppelmans
State University of Londrina (UEL) – Computer Science Department (DC) – Londrina, PR - Brazil
Email: proenca@uel.br camiel@uel.br

Mauricio Bottoli and Leonardo de Souza Mendes
State University of Campinas (UNICAMP) – Communications Department (DECOM/FEEC) – Campinas, SP - Brazil
Email: bottoli@decom.fee.unicamp.br lmendes@decom.fee.unicamp.br

Keywords: Computer network management, baseline, traffic characterization.

Abstract: This paper presents a model for automatic generation of a baseline which characterizes the traffic of network segments. The use of the baseline concept allows the manager to: identify limitations and crucial points of the network; learn about the actual status of use of the network resources; be able to gain better control of the use of network resources and to establish thresholds for the generation of more accurate and intelligent alarms, better suited to the actual characteristics of the network. Moreover, some results obtained with the practical use of the baseline in the management of network segments, are also presented. The results obtained validate the experiment and show, in practice, significant advantages in their use for network management.

1 INTRODUCTION

Computer networks are of vital importance nowadays for modern society, comparable to essential services like piped water, electricity and telephone. Extensive work has been done to improve ways to implement quality of services and traffic management along the Internet backbone (Duffield, 2001). Several existing tools and network management systems (NMS) aim to help with the network management and controls to reduce costs and improve resource utilization. However, the construction of a baseline suitable for the characteristics of each segment of a network backbone is an important task that is not usually found in the network management systems.

The Baseline can be defined as the set of basic information that shows the traffic profile in a segment of the network, through minimum and maximum thresholds about volume of traffic, quantity of errors, types of protocols and services that flow through this segment along the day. The real forecast or even an approximate one in a determined instant about the characteristics of the traffic of the segments that make up the network backbone, make the management decisions on problems that might be happening, more reliable and safer (Thottan, 2003).

The use of the baseline can help the network manager to identify limitations and control the use of resources that are critical for services that are latency-sensitive such as Voice over IP and video transport, because they can't take retransmission or even network congestion. Besides improving the resources control, its use also facilitates the planning on the network increase, for it clearly identifies the real use of resources and the critical points along the backbone, avoiding problems of performance and fault that might happen.

The use of the baseline also offers the network manager advantages related to performance management, by means of the previous knowledge of the maximum and minimum quantities of traffic in the segment along the day. This enables the establishment of more effective and functional alarms and controls, because they are using thresholds that suit the baseline, respecting the variations of traffic along the day instead of using the linear thresholds that are set based on the expertise of the human network manager (Hajji, 2003). Deviations in relation to what is being monitored real-time and what the baseline expresses must be observed and analyzed carefully, and can or can not be considered as problems. In order to do

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 149–157.

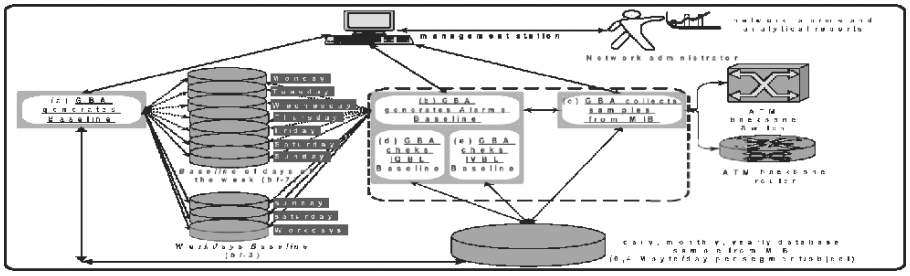


Figure 1: Operational functioning diagram for the generation of baseline and alarms.

that, the use of an alarms system integrated to the baseline and to the real-time monitoring will deal with these problems, warning the network manager when it is necessary.

As for security management, the use of the baseline can offer information related to the analysis of the users behavior, because the previous knowledge of the behavior and the traffic characteristics of a determined segment is directly related to the profile users manipulation, using this as information to prevent intrusion aspects or even network attacks, by means of the intrusion detection software (Northcutt, 2002) (Cabrera, 2001).

Another application for the baseline is related to the monitoring of a network segment which is normally performed manually by means of visual control, based on empirical knowledge with the network acquired by the manager. An example of this can be seen with the utilization of tools like GBA (Automatic Backbone Management) (GBA, 2004) and MRTG (Multi Router Traffic Grapher) (MRTG, 2004) that generate graphs with statistical analysis which consist of averages along a determined period of time about an analyzed segment or object. However, the simple use of these graphs establishes limitations for the network manager concerning to discovery and solution of problems. The limitations are caused especially by the non-automation of this task, where the monitoring of these graphs is performed visually, depending exclusively on the empirical knowledge about the functioning of the network acquired by the manager and due to the large quantity of graphs that have to be analyzed continuously. It only allows the detection of the problems and unusual situations in a reactive way.

Networks with a great number of segments turn their management more complex, considering the great quantity of graphs to be analyzed. The graphs usually present information on the volume of input and output of traffic of a certain segment, not aggregating information that could help the manager more efficiently in his decision-making with the purpose of solving problems that might be happening or that might have already happened.

Extensive work has been done in traffic characterization (Rueda, 1996), traffic measurement (Dilman, 2002) and anomaly detection (Hajji, 2003) (Thottan, 2003) (Papavassiliou, 2000), that is related to the proposal in this work. In (Rueda, 1996) is presented a survey of the main research done for traffic characterization in telecommunication networks. However these models intend to traffic modeling in a generic way, while the proposal presented in this paper intends to a traffic characterization generated from collected real data of each segment of analyzed network.

In (Hajji, 2003) is presented a proposal that is close to ours presented in this work, they proposed a baseline for automatic detection of network anomalies that uses asymptotic distribution of the difference between successive estimates of a model of network traffic. One problem that exists in this model is that it assumes that the training data is pure with no anomalies. In our case we calculate the baseline based on real data gathering from the network segment. Our baseline is generated based on statistical analyses of these data.

Thottan *et al.* (Thottan, 2003), presents a review about anomaly detection methods and a statistical signal processing technique based on abrupt change detection that uses analysis of SNMP MIB variables for anomaly detection. They use a 15s sampling frequency, and it assumes, like an open issue, that there exist some changes in MIB data that don't correspond to network anomalies. The use of an effective and real baseline can help to solve this problem for knowing the real behavior of the traffic.

Papavassiliou *et al.* (Papavassiliou, 2000), presents a tool with intend to facilitate the network management, reducing costs and minimizing the human errors. They use a similar approach to ours for the construction of baselines, when they separate workdays from weekends.

In the rest of this paper it will be presented a description about the model proposed for the

construction of the baseline, the way it was implemented and the results that show practical gains for the network management. At last, in section 3 we conclude and mention suggestions for future works.

2 BASELINE IMPLEMENTATION

The main purpose to be achieved with the construction of the baseline is the characterization of the traffic of the segment it refers to. This characterization should reflect initially the profile expected for the traffic along the day as well as other existing characteristics such as: types of protocols, types of applications, types of services. These characteristics are used to create a profile of the users. The baseline was initially developed to analyze the quantity of input and output of octates stored in the *ifInOctets* and *ifOutOctets* objects which belong to the *Interface* group of the MIB-II (RFC-1213, 1991).

The use of the GBA tool (Automatic Backbone Management) was chosen as a platform for the development of the baseline due to the great quantity of historical information related to monitoring carried out along the last years in the main network segments of UEL. The GBA was initially developed to help with the network management with ATM backbone and it performed its duty as it became a platform of learning and development, helping with the management as well as with the understanding about the networks functioning. Further information on the GBA can be found at http://proenca.uel.br/gba or in (Proença, 2001).

As for the tests and validation of the model, the data gathered by the GBA have been used since 2002 up to the present. The use of the data from the last two years was considered an important sample, characterized by periods of winter and summer vacations as well as holidays which contributed to the tests and validations of the ideas presented in this work. The analyzed data is related to the network segments with traffic TCP/IP based on Ethernet and ATM with LAN Emulation. The tests of the proposed model were carried out in three segments of the network *backbone* of UEL which are described below:

1. The first one which is called segment S_1 is responsible for interconnecting its ATM router to the other *backbone* segments; it gathers a traffic of approximately 2500 computers;
2. The second one which is called S_2 interconnects its office for undergraduate studies of academic affairs; it gathers a traffic of 50 computers;
3. The third one which is called S_3 interconnects State University of Campinas UNICAMP network to academic network at São Paulo (ANSP), it gathers a traffic of all UNICAMP (about 5000 computers) to Internet.

For the generation of the *baseline* a model was developed based on statistical analyses that we call BLGBA. The analyses were carried out for each second of the day, each day of the week. Figure 1 illustrates the operational diagram used in the implementation of the baseline, which is carried out by the GBA generated baseline module. This

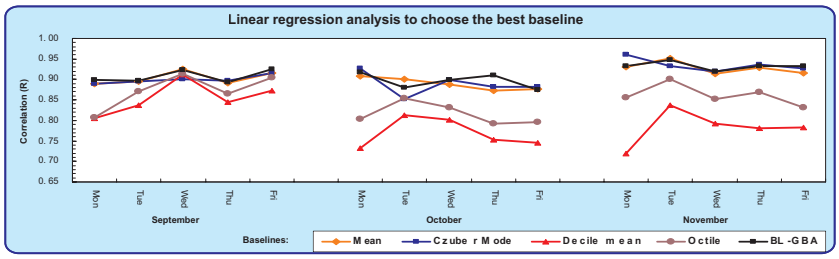


Figure 2: Linear regression analysis aiming at evaluating which is the best method for *baseline* generation.

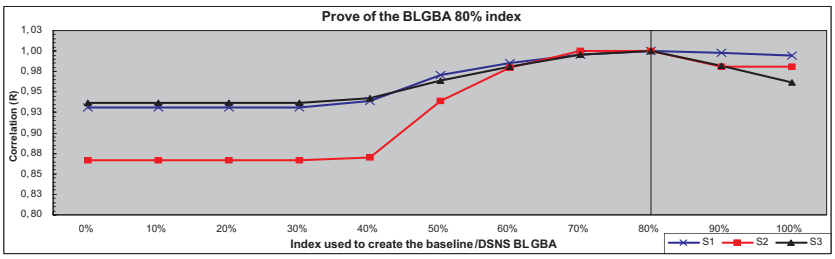


Figure 3: Linear regression analysis aiming at validating the choice for the index of 80% for the BLGBA.

module reads information from the database with data gathered daily and generates the baseline based on a period requested by the network manager.

Two types of baseline were created, one called ***bl-7*** which consists of seven baseline files, one for each day of the week, and the other one called ***bl-3*** which consists of three baseline files, one for the workdays from Monday to Friday, one for Saturday and another one for Sunday, as shown in Figure 1. The choice for generating the baseline separating the workdays of the week from Saturday and Sunday, was in order to minimize the margin of error in the final result, concerning the alterations in the volume of traffic that occur between the workdays and the other days. The results showed that it was the right choice, because the variation that was found in the volume of traffic between the workdays was of 10% and over 200% comparing workdays and weekends, as can be seen in figure 4.

The model for baseline generation proposed and presented in this work, performs statistical analysis of the collected values, respecting the exact moment of the collection, second by second for twenty-four hours, preserving the characteristics of the traffic based on the time variations along the day. For the generation of the baseline, the holidays were also excluded due to the non-use of the network on these days. Moreover, the process of baseline generation also considered faults in the collected samples which occur along the day, eliminating these faults from the calculations for the baseline generation.

The GBA makes collections at each second at the MIBs of the network equipments. Along each day, 86400 samples are expected. Problems usually occur and may affect some of these samples due to the loss of package or congesting in the network. In this case, for the generation of the baseline, the exclusion of these samples was chosen in the calculation of the baseline related to that second. This problem occurs in less than 0.05% a day, for the analyzed samples.

The processing for the baseline generation is done initially in batch aiming at its creation through data related to a pre-established period. The baseline is generated second by second for a period of days represented by N which makes up the set $\boldsymbol{n_j}$ ($\boldsymbol{j}$ = 1, 2, 3, 4, ..., N); with the daily gathering there is a set of samples of the day represented by $\boldsymbol{a_i}$ ($\boldsymbol{i}$ = 0, 1, 2, ..., 86399). Then the bi-dimensional matrix is built with 86400 lines and N columns which must be

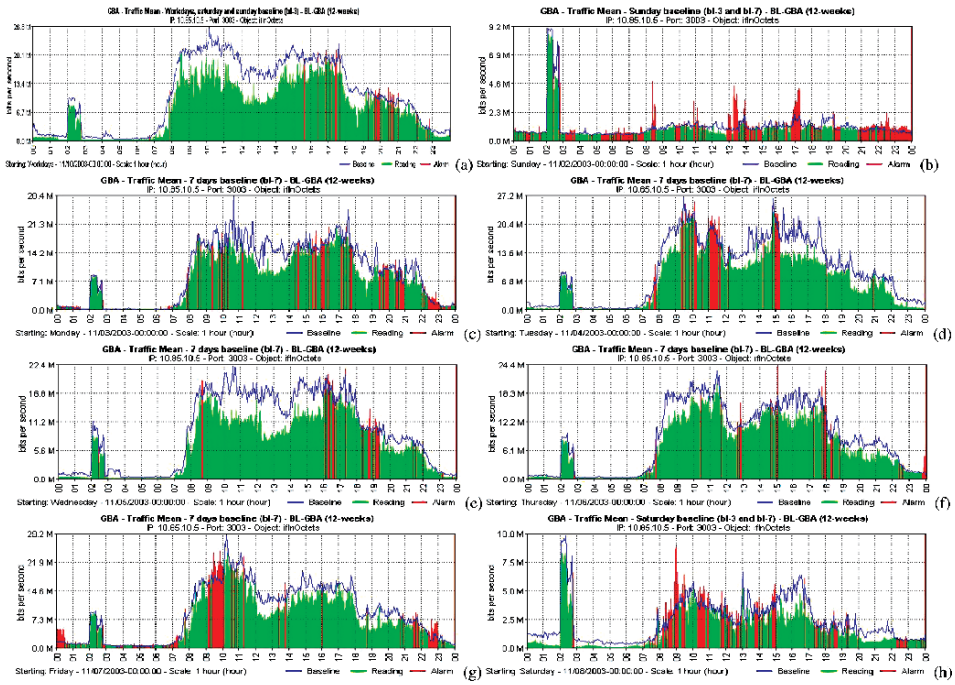


Figure 4: Baseline and the daily movement for S_1 segment analyze.

previously sorted and that will be represented by M_{ij}.

The algorithm used for the calculation of the baseline (BLGBA) is based on a variation in the calculation of *mode*, which takes the frequencies of the underlying classes as well as the frequency of the modal class into consideration. The calculation takes the distribution of the elements in frequencies, based on the difference between the greatest G_{aj} and the smallest S_{aj} element of the sample, using only 5 classes. This difference divided by five, forms the amplitude h between the classes, $h = (G_{aj} - S_{aj})/5$. Then the limits of each L_{Ck} class are obtained. They are calculated by $L_{Ck} = S_{aj} + h*k$, where *Ck* represents the k class (k = 1...5).

The proposal for the calculation of the baseline of each Bl_i second has the purpose of obtaining the element that represents 80% of the analyzed samples. The Bl_i will be defined as the greatest element inserted in class with accumulated frequency equal or greater than 80%. The purpose is to obtain the element that would be above most samples, respecting the limit of 80%. This process is used for the generation of baselines models ***bl-7*** and ***bl-3.***

The BLGBA model used for the calculation of the *baseline* was chosen after the performance of tests with other statistical models based on the *mean, octile, decile average* and on the *mode* proposed by Czuber. The choice for the BLGBA model was based on:

1. Visual analysis of graphics containing the baseline and its respective daily movement, as illustrated in figure 4;
2. Deviation analysis proposed by Bland and Altman (Bland, 1986), takes into consideration the differences between the predicted and observed movements. Such differences must lie between an interval defined by $\overline{d} \pm 2 * s$, where $\overline{d}$ is the differences mean and s is the standard deviation of these differences. With this an upper and lower limit are set where the deviation must be contained. The model that presented better adjustment was the BLGBA, with 95% of the differences in these limits;
3. Residual analysis – the model which showed less residual index between the predicted and the occurred movements was the BLGBA;
4. Linear regression (Bussab, 2003) (Papouli, 2002) between the models aimed at evaluating which one showed a better correlation coefficient between the *baseline* and the daily movement. Figure 2 shows the result of the correlation tests for the segment S_1 related to the months of September to November 2003. In this figure it is possible to notice that the BLGBA shows a better correlation coefficient between the daily movement and the *baseline*.

The choice for the element that represents 80% of the samples for the calculation of the baseline Bl_i was done empirically. Analytical tests were carried out through linear regression 00 using baseline with this value ranging between 0 and 100%, with the purpose of verifying if 80% would be the best value to be used by the BLGBA, in the calculation of the Bl_i. Figure 3 shows the correlation coefficient R between the baseline and the samples for values of choice between 0 and 100%. It is noticed that the *baseline* that uses 80%, shows a better correlation coefficient for BLGBA. These tests along with the visual analysis of the graphics with baseline and their respective daily movements showed that the value of 80% for the calculation of the Bl was the most satisfactory one

2.1 Baseline Results

The obtained results show the validity of the model for the generation of the baseline, bearing in mind the performed analyses and the comparison with the real movement that occurred. An example of that can be seen in figure 4 that illustrate in the form of a histogram, the daily movement of the segment S_1, and their respective baseline. In these figures some graphs are shown, concerning the second week in November 2003, with the *baseline* in blue and the real movement that occurred on the day in green. We came to the following conclusions with the results shown in figures 4:

1. Clear peaks of traffic in the baseline everyday between 0:30 and 4:00 o'clock in the segment S_1 that are related to the backup performed in this period in the network server;
2. The profile of traffic for the workdays, figures 4 (a), generated by the ***bl-3*** model and 4 (c), (d), (e), (f) and (g), generated by the ***bl-7*** model, is quite similar with a strong time dependence along the day which, in this case, is related to the working day hours of the university where the tests were performed. In the case of Saturdays and Sundays, the baseline generated for these days are exactly the same for ***bl-3*** and ***bl-7*** models, figures 4 (b), (h) shows this results;

Table 1: Variation of the baseline from January 2003 to January 2004, for segment S_1.

% of growth of the baseline/DSNS comparede with the previous month													
	Jan/03	Fev/03	Mar/03	Apr/03	May/03	Jun/03	Jul/03	Aug/03	Sep/03	Oct/03	Nov/03	Dec/03	Jan/04
IVBL	1,10%	1,51%	5,38%	0,07%	8,66%	2,94%	5,83%	6,38%	4,95%	4,12%	2,78%	2,89%	3,02%

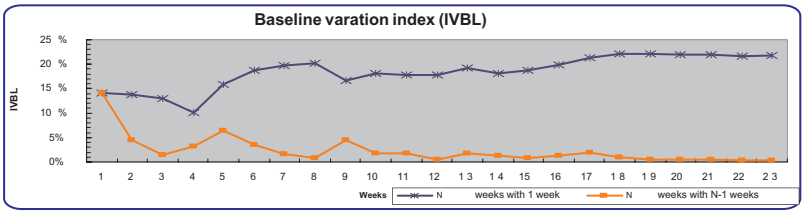


Figure 5: % of variation of the baseline of *n* weeks compared to the (n – 1) weeks and 1 week.

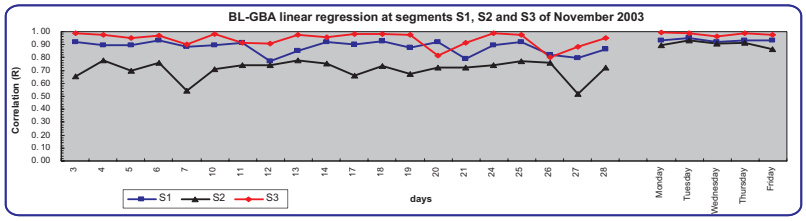


Figure 6: Analysis of the BLGBA by linear regression of November 2003.

3. Not only the baseline generated for the workdays ***bl-3*** but also the one generated for all the days of the week ***bl-7***, showed to be suitable for the characterization of the traffic. The ***bl-7*** is a model of baseline to be used in cases in which there is the need to respect individual particularities which occur in each day of the week, such as backup days, whereas the ***bl-3*** is the most suitable for the cases where this is not necessary, that is, all the workdays can be dealt with in a single baseline, leaving the decision on what model to be used to the network manager's;
4. Periods in which the traffic of the day becomes higher than the baseline. In this case, its color is changed from green to red, which means a peak of traffic above the baseline, and this could or could not be interpreted as an alarm;
5. The generated baselines fulfill their main objective which is the characterization of the traffic in the analyzed segments;
6. The baseline is influenced by time factors which, in this case, are related to the working day that starts at 8:00 a.m. and finishes at 10:00 p.m.
7. The baselines presented in figures 4 were generated by a 12 week sample collection of real data in segment S_1. Our studies have demonstrated that for segments with a lot of aggregate traffic as in S_1 and S_3, 12 weeks is necessary for a baseline formation.

Unfortunately, due to the limited quantity of information that is presented in this article, it is not possible to show other figures which corroborate what was presented in this work. Nevertheless, at the address http://gba.uel.br/blgba more information and results obtained through this work can be found.

2.2 Baseline Evaluation

We created an index with the purpose of evaluating the coefficient of variation of the baseline of one month in relation to the other. This index is called Index of Variation of the Baseline (IVBL). The IVBL is calculated based on the difference between one baseline and the other, as shown in equation (1). With the IVBL it was possible to conclude that there is usually a positive variation in the volume of traffic from one month to the other, showing that despite being small, there is a tendency of growth in the volume of traffic in the analyzed segments.

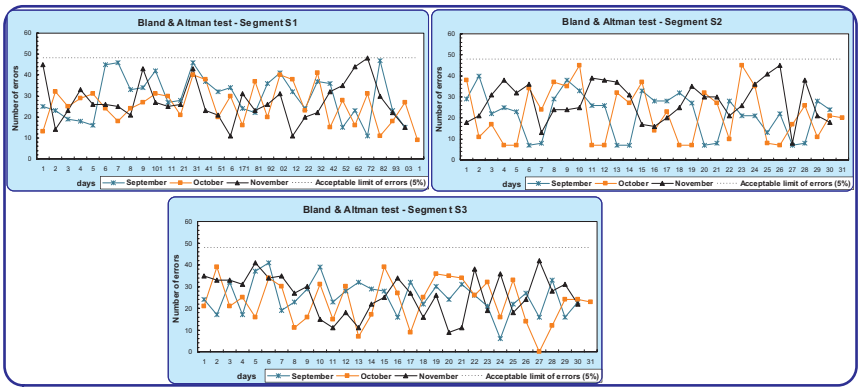


Figure 7: Bland & Altman test from September to November 2003 for segments S_1, S_2, and S_3.

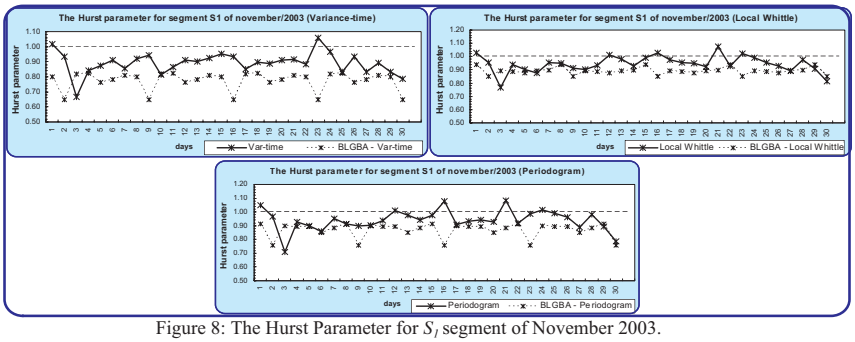


Figure 8: The Hurst Parameter for S_1 segment of November 2003.

Table 1 shows the percentage of growth in the segment $\boldsymbol{S_1}$ from the network of UEL, from January 2003 to January 2004. In the other analyzed segments, a small percentage of growth was also observed.

$$IVBL = \left(\sum_{i=1}^{86400} BL'_i - BL''_i \right) \Big/ 86400 \qquad (1)$$

Where *IVBL* = variation index of a *baseline* in relation to another

The IVBL was also used to calculate the variation of a baseline generated from $\boldsymbol{n}$ weeks and compared to a baseline of $(\boldsymbol{n} - 1)$ weeks, and in the comparison between the baseline of 1 week with the baseline of $\boldsymbol{n}$ weeks. These calculations using weekly baselines were carried out with the purpose of evaluating and demonstrating the minimum quantity of samples necessary for the formation of the baseline. Initially it was concluded empirically that it would be necessary 4 to 12 weeks for the formation of the baseline. With the comparison of the baseline of $\boldsymbol{n}$ weeks with the one of $(\boldsymbol{n} - 1)$, during 24 weeks, it was observed that the percentage of variation tends to stabilize from the 12th week on, and not being significant for the formation of the baseline. And when a baseline of 1 week was established and a comparison was carried out for 24 weeks, it was also noticed that, from the 12 th week on, the percentage of variation tends to stabilize around 20%, showing no more significant variations that could be added to the baseline from this point on. The figure 5 shows the results of these comparisons.

Besides the visual evaluation of the results, other analytical tests have been carried out aiming to evaluate the reliability of the baseline generated by the BLGBA in relation to the real movement. The tests were carried out from January to November of 2003, below is presented a synthesis of the results:

I. Linear Regression (Bussab, 2003) (Papoulis, 2002): Figure 6 presents the results of the linear regression for the segments $\boldsymbol{S_1}$, $\boldsymbol{S_2}$, and $\boldsymbol{S_3}$ for all workdays of November of 2003. The results demonstrate a high correlation and adjustment between the movement that

occurred those days in relation to their baseline;

II. Test purposed by Bland & Altman (Bland, 1986): Refer to the deviations analysis that occur between the baseline and the real movement. 95% of the deviations/errors observed during all days from September to November 2003, in segments S_1, S_2, and S_3, are between the required limits of $\overline{d} \pm 2 * s$, where $\overline{d}$ is the mean and s is the standard deviation of the differences between the baseline and the real movement, as shown in figure 7. In the other months of the year the results had also confirm the reliability of the model, keeping 95% of the cases inside the limits established of $\overline{d} \pm 2 * s$;

III. Hurst parameter (H): Tests carried out with the real movement and the baseline generated by the BLGBA, using the statistical methods Variance-time, Local Whittle and Periodogram (Leland, 1994) generate the hurst parameter *H*. The analysis confirms that the traffic is self-similar and the baseline is also self-similar, however presenting a lower hurst parameter. Figure 8 illustrates an example of these calculations for S_1 segment during November 2003. In most of the cases, these tests also allow us to notice that in segments with lower number of computers like S_2, the hurst parameter presents a lower rate between 0.6 and 0.7, in segments with great aggregated traffic like the S_1, and S_3 it presents a rate between 0.8 and 1.0. The Hurst parameter evaluation was made using the samples collected second by second with the GBA tool. Calculations were made for each day between 8:00 and 18:00 hours, the period when the traffic is more similar to a stationary stochastic process. Its utilization makes possible the evaluation of the baseline quality in segments of different burstiness. Indicating that the greater the burstiness of the segment, the bigger the Hurst parameter and the better the characterization showed by the baseline. And the lower the burstiness of the segment, the smaller the Hurst parameter and worse the results shown by the baseline. These results are corroborated by the other tests utilized to validate the baseline that also indicate an increase of the baseline quality in segments with a higher burstiness.

3 CONCLUSION

This work presented a contribution related to the automatic generation of baseline for network segments, which constitutes itself into an important mechanism for the characterization of the traffic of the analyzed segment, through thresholds that reflect the real expectation of the volume of traffic respecting the time characteristics along the day and the week. This enables the network manager to identify the limitations and the crucial points in the network, control the use of the network resources, establish the real use of the resources, besides contributing to the planning of the needs and demands along the backbone.

The use of an alarms system integrated to the baseline as well as with the monitoring performed real time by the GBA, figure 1 (b) and (c), can make possible for the network manager to be informed through messages, at the exact moment a difference related to the expected traffic and the baseline, was found out. This possibility is fundamental for the segments or crucial points of the networks that demand perfect control and pro-active management in order to avoid the unavailability of the services rendered.

The use of graphs such as the ones shown in figures 4 with information about the baseline and about the daily movement, makes a better control over the segments possible.

It could be noticed that the behavior of the traffic of the Ethernet networks is random, self-similar and extremely influenced by the quantity of bursts, which intensify as the number of hosts connected to the segment increase, as shown in (Leland, 1994). It also showed that the model chosen for the characterization of the baseline, presented in this work, is viable for the characterization of the traffic in backbone segments that concentrate the traffic of a great number of hosts, as shown in the examples of section 2.

Tests were also realized with baselines from other MIB objects, like ipInReceives, icmpInMsgs, udpInDatagrams. The results have been satisfactory and demonstrated that the BLGBA model can be used for other MIB objects.

Besides the tests performed at the networks of UEL and be initiate in the Communications Department of the Electric Engineering Faculty of UNICAMP, which results validating the model presented in this work, tests with different types of networks, such as factories, large providers and industries shall be performed, aiming to evaluate

and perfect the model proposed for generation of baseline.

Another future work being developed refers to the creation of a multiparametric model for alarms generation aiming to aid the security, performance and fault management, using a set of some monitored objects baseline, such as IP, TCP, UDP and ICMP packet traffic, traffic volume in bytes and number of errors. The model consists in the utilization of a baseline set, information about possible network anomalies and rules for alarm generation based on thresholds in differentiated levels, which would indicate specific conditions to customizable problems to the network. A creation of an efficient mechanism of anomaly detection and alarm generation is expected.

REFERENCES

Duffield, N.G.; Grossglauser, M. (2001, June) Trajectory sampling for direct traffic observation; Networking, IEEE/ACM Transactions on, Volume: 9, Issue: 3, Pages: 280 - 292.

Cabrera, J.B.D.; Lewis, L.; Xinzhou Qin; Wenke Lee; Prasanth, R.K.; Ravichandran, B.; Mehra, R.K. (2001, May); Proactive detection of distributed denial of service attacks using MIB traffic variables-a feasibility study, Integrated Network Management Proceedings, IEEE/IFIP International Symposium on , Pages: 609 - 622.

Northcutt, Stephen, Novak Judy. (2002) Network Intrusion Detection, Third Edition, New Riders.

GBA, Ferramenta para Auxílio no Gerenciamento Backbone Automatizado, Retrieved 03/05/2004 from http://proenca.uel.br/gba/.

MRTG, The Multi Router Traffic Grapher, Retrieved 03/05/2004 from http://people.ee.ethz.ch/~oetiker/webtools/mrtg/.

Rueda, A.; Kinsner (1996, May); A survey of traffic characterization techniques in telecommunication networks, Electrical and Computer Engineering, Canadian Conference on, Vol.2, Pages:830-833.

Dilman, M.; Raz, D. (2002, May) Efficient reactive monitoring Selected Areas in Communications, IEEE Journal on, Vol. 20, Iss. 4, Pages: 668-676.

Hajji, H. (2003, May); Baselining network traffic and online faults detection; Communications, ICC '03. IEEE International Conference on, Vol. 1, 11- Pages: 301 - 308.

Thottan, M.; Chuanyi Ji (2003, Aug); Anomaly detection in IP networks, Signal Processing, IEEE Transactions on Volume:51, Issue:8, Pages: 2191-2204.

Papavassiliou, S.; Pace, M.; Zawadzki, A.; Ho, L. (2000, June); Implementing enhanced network maintenance for transaction access services: tools and applications, Communications, 2000. ICC 2000. IEEE International Conference on, Volume: 1, 18-22, Pages: 211-215 vol. 1.

Proença, Mario Lemes, Jr. (2001, September) "Uma Experiência de Gerenciamento de Rede com Backbone ATM através da Ferramenta GBA", Artigo publicado no congresso, XIX Simpósio Brasileiro de Telecomunicações – SBrT 2001, Fortaleza 03-06/09/2001.

RFC-1213, INTERNET ENGINEERING TASK FORCE (IETF) (1991, March) Management Information Base for Network Management of TCP/IP-based internets: MIB-II.

Bland J. Martin and Altman Douglas G. (1986), Statistical Methods For Assessing Agreement Between Two Methods of Clinical Measurement, The LANCET i: 307-310, February 8, 1986.

Bussab, Wilton O.; Morettin Pedro A. (2003) Estatística Básica, Editora Saraiva, 5a edição.

Papoulis, Athanasios, Pillai S. Unnikrishna. (2002) Probability, Random Variables and Stochastic Processes, Fourth Edition, McGraw-Hill.

Leland Will E., Taqqu M. S., Willinger W., Wilson D. V., (1994) On the Self-Similar Nature of Ethernet Traffic (Extended Version), IEEE/ACM Transactions on Networking, vol. 2, No. 1, February 1994.

NETWORK-BASED INTRUSION DETECTION SYSTEMS EVALUATION THROUGH A SHORT TERM EXPERIMENTAL SCRIPT

Leonardo Lemes Fagundes and Luciano Paschoal Gaspary
Programa Interdisciplinar de Pós-Graduação em Computação Aplicada,Universidade do Vale do Rio dos Sinos
Av. Unisinos 950 – 93.022-000 – São Leopoldo, Brazil
Email: leonardo@exatas.unisinos.br, paschoal@exatas.unisinos.br

Keywords: Security, intrusion detection systems, evaluation.

Abstract: Intrusion Detection Systems (IDSs) have become an essential component to improve security in networked environments. The increasing set of available IDSs has stimulated research projects that investigate means to assess them and to find out their strengths and limitations (in order to improve the IDSs themselves) and to assist the security manager in selecting the product that best suits specific requirements. Current approaches to do that (a) require the accomplishment of complex procedures that take too much time to be executed, (b) do not provide any systematic way of executing them, and (c) require, in general, specific knowledge of IDSs internal structure to be applied. In this paper we address these limitations by proposing a script to evaluate network-based IDSs regarding their detection capability, scalability and false positive rate. Two Intrusion Detection Systems, Snort and Firestorm, have been assessed to validate our approach.

1 INTRODUCTION

With the large scale use of Internet, the number of attacks against all kinds of organizations has increased considerably. Through the exploration of different types of vulnerabilities such as configuration flaws, implementation flaws and improper use of available resources, a universe of possible attacks emerge. Examples of such attacks go from port scanning, denial of service, connection hijacking to more sophisticated attacks, such as distributed denial of service, insertion and evasion. Aiming at minimizing an intruder's chances to obtain success in his/her activities, several protection mechanisms are used. Among these mechanisms are cryptography, digital certification, public key infrastructures, firewalls, authentication protocols and intrusion detection systems.

The IDSs represent an important monitoring technique, whose main function is to detect malicious actions, such as attack attempts and illegal access to information. There are several intrusion detection systems available in the market. Among these IDSs some that stand out are Snort (Roesch, 1999), Bro (Paxson, 1999), NFR (Network Flight Recorder) (NFR, 2001), Firestorm (Firestorm, 2001) and RealSecure (ISS, 1999). Considering the increasing number of IDSs, the identification of their strengths and limitations is essential, not only to stimulate specific research niches in the area (to improve the IDSs), but to assist the security managers in their grueling task of selecting the most appropriate system for the environment used as well.

Several approaches have been proposed to assess intrusion detection systems (Puketza et al., 1997; Lippmann et al., 1999; Alessandri, 2000; Barber, 2001). Those approaches, however, don't describe a systematic way of executing the procedures, presenting a series of exhausting activities that take several weeks and demand specific knowledge from the users, such as the internal structure of IDSs (which is not possible in case of proprietary IDSs).

This paper presents an alternative approach to evaluate network-based IDSs by presenting a script with a set of systematic procedures, which can be done in a short period of time and that do not require previous knowledge of the detection tools to be assessed. Despite the fact that the results of the evaluation don't present as many details as some approaches just mentioned, they can be easily carried out (without requiring highly specialized human resources and materials). The tests here evaluate the following IDS characteristics: detection capability, scalability and false positive rates. These characteristics were chosen for they are the most

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 159–165.

representative to an organization decision process when choosing an IDS. It is also worth mentioning that our approach has been proposed to comparatively access two or more intrusion detection systems (and not to measure the individual capabilities of a single system). The paper is organized as follows: section 2 presents some related work. Section 3 describes the script proposed to conduct the assessment. The results obtained with Snort and Firestorm IDSs are presented in section 4. Section 5 concludes the paper with some final remarks and presents perspectives for future work.

2 RELATED WORK

This section presents the main approaches developed up to this date to assess intrusion detection systems. The criteria applied in this comparison were: type of assessment, nature of background traffic generated to perform the experiments and requirement of test settings.

Regarding the type of assessment, it has been observed that most of the approaches assess only the IDSs' signatures bases (Puketza et al., 1997; Lippmann et al., 2000, Barber, 2001) which, in addition to being exhaustive work considering the size of these bases, generates a valid result only for a short time period, because signatures are developed by IDSs' creators or even by users quite frequently. Therefore, in order to consider the results of these approaches reliable, it is necessary to run all the experiments every time a new signature is released. On the other hand, approaches such as the ones proposed in this paper and in (Alessandri, 2000), instead of testing the signature base, actually test the IDS's detection capabilities. By using them, experiments must be re-run only when new detection functionalities are added to the system itself.

The representation of background traffic is a fundamental feature in the assessment of IDSs, because it interferes directly in the results of some tests, such as the ones on false positive rates and scalability. There are approaches like (Lippmann et al., 1998) and (Lippmann et al., 1999) that do not describe how background traffic is composed. This leads to the objection of results, especially in the assessment of false positives, because there is no way to assure if there are attacks inserted on this traffic, nor there are ways to identify which reasons might have led the IDSs to generate such results.

Except for the methodology proposed by (Alessandri, 2000), all the others require some sort of test setting to run the experiments. Approaches such as (Puketza et al., 1997) and (Lippmann et al., 1999) require complex test settings, with dozens of stations (attackers, victims, evaluated systems, traffic generators and traffic collectors), various interconnectivity equipments (hub, switch and routers), and even firewalls. These requirements in a test setting often result in an impracticable choice, due to the fact that they demand an extended time period and a dedicated environment up until completion of tests. Furthermore, the use of firewalls prevents several attacks from being captured by the IDSs, since they are blocked before reaching the company's internal network. The use of firewalls is extremely crucial for any business and must be part of every study which aims at assessing security infrastructure. However, considering this specific purpose, it is a factor that limits the assessment process.

As a general rule, what is observed in the approaches quoted is that they lack a proposal which could be applied by organizations security staff. In order to accomplish that, it is necessary to develop an approach that presents well defined procedures, that can be easily achievable, and that can fully reflect the reality of the criteria assessed. The approaches referenced here fail in these aspects. In addition to not providing adequate documentation on how some important tests were conducted, some of these proposals have not yet been properly validated, or do not offer the necessary means to be applied.

3 EVALUATION SCRIPT

The script is composed by five steps: selection of attacks, selection of tools, generation of traffic settings for evaluation, assembly of evaluation environment and IDSs analysis.

3.1 Selection ofAttacks

The goal is to select a set of attacks that present distinct technical characteristics amongst them. Instead of simply gathering a set of attacks, we propose to select, by the end of this phase, attacks whose detection is possible through different existent mechanisms in an IDS. For instance, for an IDS to be capable to detect an insertion attack, the *URL Encoding,* it needs more than the capability of analyzing an HTTP packet, because a content decoding mechanism of the packet header is also necessary. Similarly, to detect denial of service attacks such as the *teardrop* a mechanism capable to rebuild fragmented IP packets is required. Thus, the attacks selected in this step present a unique set of characteristics that allows the evaluation of different IDSs detection capabilities and not only the signature database.

				HTTP			IP					TCP										ICMP		UDP
	Multiple packets	Does not establish a connection	Establishes a connection	Requisition line	Requisition encoding	Requisition size	Fragment off-set	IP packet with an enabled DF bit	Fragment identification	Raw IP packet	Fragmentation control flags	Options (NOP, MSS, Windows, ...)	Field: service type	Sequence number	TCT initial window	Packet with SYN flag	Packet with FIN flag	Packet with ACK flag	Packet with URG flag	Packet with PUSH flag	Packet with all flags disabled	ICMP packet	ICPM error message size	0 bytes UDP packet
TCP Connect	X		X																					
Syn scan	X	X														X								
Ack Scan	X	X																X						
Window Scan	X	X																X						
Fin scan	X	X															X							
UDP Scan	X	X																						X
Null Scan	X	X																			X			
Xmas	X	X															X	X	X	X				
TCP Ping	X	X																X						
TCP Fragmentation	X	X					X		X		X					X	X	X	X	X				
IP Scanning	X	X								X														
Fingerprinting	X	X						X				X	X	X	X		X	X					X	
Ident Reverse TCP	X		X															X						

Figure 1: Technical description of the initial attack setting proposed.

Figure 1 shows the characteristics explored by the set of possible attacks to be used in the IDSs evaluation (in columns). In the lines all attacks are listed. Due to space limitation, the figure illustrates only the port scanning attack, but evasion, insertion and denial of service attacks were also considered. The attacks used in the assessment are those that explore combinations of different characteristics (in bold). To ease the problem that the IDS does not have the signature of the selected attack and to avoid reaching a conclusion that the IDS is not capable to detect the attacks, we always selected two attacks that explore the same characteristics: one past and one recent. This approach allows reducing the initial attack setting. We assume, for instance, that if an IDS is capable to detect an *Ident Reverse TCP* attack, it should also be capable to detect a *TCPConnect* (that explores the same characteristics); it only needs to be configured with appropriate signatures to do so.

3.2 Selection of Tools

This step is dedicated to the obtention of tools that allow reproducing the attacks selected in the previous stage. This task can be accomplished in a short time period, since the tools applied are easy to find and use. For instance, to reproduce the port scannings listed in figure 1 the tool Nmap 2.54 (GNU/Linux) could be used.

3.3 Generation of the Evaluation Trafic Settings

The evaluation set up is formed by the selected attacks and by the background traffic (necessary for the scalability analysis). Following, we present the description proposed in the script (a) to store the attack traffic and (b) to generate the background traffic.

3.3.1 Gathering of the Attack Traffic

In order to avoid the manipulation of each attack tool every time that the several tests (presented in section 3.5) have to be performed, we suggest the previous gathering and storage of the attack traffic. In order to do that, an environment such as the one illustrated in figure 2a must be setup. In the *Attacker* station all attack tools selected in the tools selection step are installed, while the *Victim* station has all services to be attacked installed. Finally, the *Sniffer* station collects the traffic (using a tool such as tcpdump) generated by both the attack tools and the attacked station (when, in some level, it presents a reaction to them). So, for each attack to be collected and stored, the following sequence of steps is suggested: (a) start up the tcpdump, (b) execute the attack, (c) stop the tcpdump and (d) store the traffic.

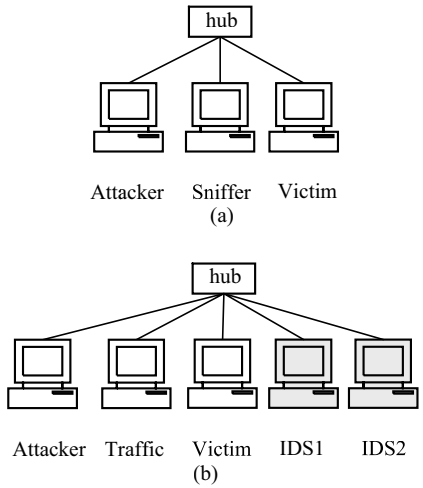


Figure 2: Network environment (a) to collect and store the traffic of attacks and (b) to evaluate the IDSs.

3.3.2 Generation of Background Traffic

The background traffic is necessary in order to analyze the IDSs scalability. Our group decided to use uniform artificial background traffic in the analysis because when using real traffic, due to the throughput rate oscillation and the alternation of the applications used, it would be difficult to identify, for instance, at which rate the IDS starts to to discard packets. Besides, the use of real traffic (unless the traffic was known in full detail), could lead to non-foreseen alarms (for instance, if there were any attacks inserted in that traffic), what would generate noise in the evaluation in course.

So we propose to use a background traffic composed by 256 bytes UDP packets. This traffic must be reproduced at different rates (e.g. 4, 6, 8, 10 and 12 Mbps). An appropriate tool to do so is `udp_generator`, developed at LAND/Federal University of Rio de Janeiro, since it is easy to manage and, for that, makes it unnecessary to store the traffic for subsequent reproduction.

3.4 Assembly of the Assessment Environment

The evaluation environment should be composed by the IDSs to be evaluated, the target stations (*Victims*) that will undergo the attacks, a station to repro-duce the attack traffic (*Attacker*) and another to reproduce the background traffic (*Traffic*) during the scalability tests. Figure 2b shows the environment used in our evaluation, while the results are presented in section 4.

The number of victim stations and the operating system installed in those stations may vary according to the attacks selected to compose the evaluation setting. For instance, if the evaluation setting is comprised of attacks to Solaris and Windows 2000 Server stations, the network environment represented in figure 2b should use two additional victim stations, in which those systems should be properly installed and configured. Also, the amount of IDSs evaluated may vary. As consequence, the number of stations to host these systems can also be larger.

3.5 IDSs Analysis

As already mentioned, the script proposed evaluates the following IDSs characteristics: detection capability, scalability and false positive rates generated by these systems. Detection capability is the test that allows identifying the IDSs detection strengths, in other words, which types of attacks the system is able to detect. The scalability test allows identifying at which transmission rate the IDS starts to discard packets. The false positive rate shows the tendency of the IDS to generate false alarms. It occurs when normal traffic is erroneously considered an attack or when an attack takes place but the IDS generates alarms informing of other attacks instead of the one going on.

3.5.1 Detection Capability

To evaluate the IDSs detection capability the following sequence of steps is suggested: (a) clean

the log files and begin the IDSs log service, (b) reproduce, one by one, the traffic collected from the attacks (see section 3.3.1), (c) stop the log service, (d) save the generated files, and (e) count and identify the detected and non detected attacks (from the log analysis). The reproduction of the attacks will be triggered from the Attacker station, at a low rate (so the IDS does not discard packets), using a tool such as `tcpreplay`.

3.5.2 Scalablity

For the scalability analysis results to be reliable, only the attacks detected by the IDSs in the previous analysis are reproduced. For example, if in an evaluation A five attacks are identified and an evaluation B detects only three, the scalability analysis is going to consider five and three attacks, respectively.

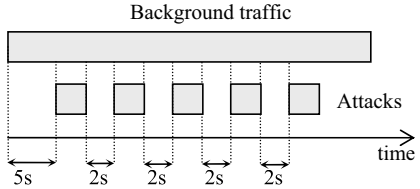


Figure 3: Traffic generation and reproduction sequence used for scalability analysis.

Figure 3 represents the relationship between the attack traffic and the background traffic (uniform, generated at a constant rate), reproduced simultaneously in this analysis. Each type of attack considered in the evaluation (in our case, denial of service, evasion, insertion and port scanning) must be executed in a different series of tests. Figure 3, for example, refers to the scalability analysis of an IDS under the denial of service attacks. In this case, five attacks (Smurf, UDP Storm, Syn Flood, Teardrop and ICMP Fragmentation) are being reproduced in parallel with the background traffic (a). The attacks are executed one by one with a two second interval between them. The generation of background traffic starts five seconds before the first attack is reproduced and is only stopped after all attacks have been transmitted. So, as soon as that sequence has been terminated, (b) the *log* system must be stopped, (c) the number of alerts generated by the IDS should be stored (for later account), (d) the *log* system must be restarted and, soon afterwards, (e) the next type of attack must be reproduced. After every possible type of attack has been reproduced for each IDS, (f) the background traffic transmission rate must be increased and the procedure described must be repeated (a).

3.5.3 False Positive Rates

False positive is every alarm that indicates that an attack is happening, when actually another kind of activity is taking place. For example, when a support user executes a ping command for a server and the IDS stores this event as an attack. Another example occurs when the network is actually under one type of attack (e.g. *UDPstorm*), but the IDS generates alarms not only for this attack but for other type of attacks, that are not happening at the moment (e.g. *ICMP fragmentation*).

Our proposal to identify the false positive is based on the analysis of the *logs* generated in the detection capability tests. After a set of attacks, the IDS *log* stores the alarms. The security manager can easily sort the alarms that identify actual attacks and false positive alarms. The ratio between the number of additional alarms over the total of alarms generated for a set of attacks represents an important indicator of the IDS tendency to generate false positives. For that, the *log* files generated by the detection capability test should be checked again.

4 CASE STUDY

This section presents the results achieved by two IDSs submitted to our experimental script described in the previous section. The IDSs used in this case study were Snort 1.83 (Roesch, 1999) and Firestorm 0.4.6 (Firestorm, 2001), both available under GNU GPL version 2 license.

4.1 Detection Capability

The detection capability analysis was carried out based on the attacks mentioned in section 3.1. This analysis was made simultaneously with the two chosen IDSs. The results achieved by Snort and by Firestorm are shown in Figure 4. It is important to emphasize that, although the sequence of tests described in section 3.5.1 has been repeated ten times, the results were always the same (statistical variance equals zero).

The results demonstrate that Snort is a tool capable to detect insertion, evasion, port scanning and denial of service attacks very efficiently. Firestorm, on the other hand, did not present an efficient mechanism to decode HTTP requests.

4.2 Scalability

To evaluate the scalability, the IDSs were submitted to the procedure described in section 3.5.2, with

background traffic of 4, 6 and 8 Mbps. The tests were carried out ten times and presented a 2.5% variance. It was observed that, at a transmission rate of 4 Mbps, the IDSs don't present packet loss. So, the number of alerts generated (including false positive) corresponds to the maximum possible for the set of attacks being analyzed.

	Snort	Firestorm
Evasion		
Method Matching	X	X
Session Splicing	X	X
Insertion		
Long URL	X	X
Self Reference	X	X
URL Encoding	X	
Port scanning		
UDP Scan	X	
Xmas	X	X
TCP Fragmentation	X	X
IP Protocol Sweeping	X	
Fingerprinting	X	X
Ident Reverse TCP	X	X
Denial of Service		
Smurf	X	X
UDP Storm	X	X
Syn Flood	X	X
Teardrop	X	X
ICPM Fragmentation	X	X

Obs.: The "x" means that the respective IDS detected the attack. If the space is blank it means that the IDS did not detect the attack.

Figure 4: Detection capability analysis results.

	Evasion	Insertion	Port Scanning	Denial of Service
4 Mbps				
Snort	100%	100%	100%	100%
Firestorm	100%	100%	100%	100%
6 Mbps				
Snort	98,81%	97,86%	99,32%	99,83%
Firestorm	97,56%	95,18%	99,57%	99,36%
8 Mbps				
Snort	89,99%	86,10%	94,85%	94,41%
Firestorm	86,04%	83,42%	90,49%	92,24%

Obs.: The value in each cell is obtained by dividing the number of *logs* stored (including false positive) by the number of maximum alarms expected.

Figure 5: Scalability analysis results.

In figure 5, the results obtained with the IDS evaluation are presented. As it can be observed, Snort is slightly superior compared to Firestorm regarding the scalability analysis as to all attack types considered. However, both IDSs can fail to detect attacks even at a low transmission rate (8 Mbps).

4.3 False Positive Rate

The false positive rate analysis, as described in section 3.5.3, is done by using the *log* file data generated by the IDS when executing a detection capability analysis. The sequence of tests described in section 3.5.3 was carried out ten times, always obtaining the same results (statistical variance equals zero).

Figure 6 shows the false positive rates generated by Snort and by Firestorm. The high values indicate that the IDSs has an imprecise set of signatures, what causes the mistaken generation of alarms. The table demonstrates that this matter is more critical in Firestorm, although the results presented for Snort are not encouraging as well.

	Evasion	Insertion	Port Scanning	Denial of Service
Snort	36,73%	29,97%	4,71%	4,63%
Firestorm	41,32%	42,97%	12,93%	7,47%

Obs.: The value in each cell is obtained by dividing the number of additional alarms (false positives) by the total of alarms generated.

Figure 6: False positive rate analysis.

5 CONCLUSIONS AND FUTURE WORK

Nowadays the main approaches regarding IDS assessment use a large amount of reproduced attacks, because it is believed that this allows the evaluation to be more detailed. However, there aren't any preexistent criteria for the selection of the attacks. Therefore, many of them explore the same characteristics making a wide and detailed evaluation of the IDS strengths impossible. Besides, the selection of attacks leads to extremely

exhausting experiments given the amount of attacks that they analyze. The script proposed in this paper describes a method for attack selection, in which the setting used in the IDSs analysis is composed only by attacks that present unique characteristics. Through this selection, described in section 3.1, the initial attack setting is reduced by approximately 50%.

Regarding the IDSs scalability evaluation, it can be said that the sustained capacity of the system being evaluated is very clear, even though the script is based on a uniform traffic rate. Another aspect is that even when submitted to low traffic, the IDSs begin to discard packets (compromising the detection process). It still remains to be done an extended scalability evaluation of those IDSs for rates higher than 10 Mbps (ex: up to 100 Mbps).

The evaluation of false positive rates generated, as proposed in this paper, is influenced by the power of description languages to describe signatures and by the precision of the network manager when specifying them. An additional mechanism for this analysis, that takes into account the typical background traffic found in the organization where the IDS is going to be used, requires further investigation.

Future work include (a) the extension of the evaluation setting, through the selection of other types of attacks, (b) the investigation of procedures to evaluate other criteria (ex: capacity to handle concurrent attacks) and (c) the development of a tool to assist the execution of the script proposed.

REFERENCES

Alessandri, D. (2000). Using rule-based activity to evaluate intrusion – detection systems. In *Third International Workshop on Recent Advances in Intrusion Detection (RAID)*, pages 183-196.

Barber, R. (2001). The evolution of intrusion detection systems – the next step. *Computer & Security*, 20(2):132-145.

Firestorm (2001). *Firestorm network intrusion detection system Homepage*. http://www.scaramanga.com.uk/.

ISS (1999). *Real Secure Systems Inc. Homepage*. http://iss.net.

Lippmann, R., Fried, David J., Graf, I., Haines, Joshua W., Kendall, Kristopher R., McClung, D., Weber, D., Webster, Seth E., Wyschogrod, D., Cunningham, Robert K. and Zissman, M. (1998). Evaluating intrusion detection systems: The 1998 DARPA off-line intrusion detection evaluation. In proceedings of the DARPA Information Survivability Conference and Exposition (DISCEX) 2000, IEEE Computer Society Press, Los Alaminos, CA.

Lippmann, R., Haines, D., Fried, D. J., Das, K. J., and Korba, J. (1999). Evaluating intrusion detection systems the 1999 darpa off-line intrusion detection evaluation. *Computer Networks*, 34 (4):579-595.

NFR (2001). *Network Flight Recorder, Inc. Homepage*. http://www.nfr.com/.

Paxson, V. (1999). Bro a system for detecting network intruders in real-time. *Computer Networks*, 31(23-24) 2435-2463.

Puketza, N., Chung, M., Olsson, R. A., and Mukherjee, B. (1997). A software platform for testing intrusion detection systems. *IEEE Software*, 14(5):43-51.

Roesch, M. (1999). Snort – lightweight intrusion detection for networks. In *USENIX LISA Conference*.

A SINGLE SIGN-ON PROTOCOL FOR DISTRIBUTED WEB APPLICATIONS BASED ON STANDARD INTERNET MECHANISMS

Julian Gantner, Andreas Geyer-Schulz and Anke Thede

Information Services and Electronic Markets
Universität Karlsruhe (TH), 76128 Karlsruhe, Germany
Email: {gantner,geyer-schulz,thede}@em.uni-karlsruhe.de

Keywords: single sign-on, cookies, Web authentication, Web services, cross-domain

Abstract: Growing e-commerce and personalized Web sites require users to set up many different personal accounts. Personal data has to be entered many times and each user has to memorize different username and password combinations. This reduces system security as users tend to either use passwords that are very easy to guess, or they write them down, or they use the same password for many different accounts. It also increases the cost of the administration of the user accounts.

We propose a protocol for a single sign-on system that allows users to visit multiple internet applications having to login only once. The system is based on standard internet mechanisms. It is composed of different servers that provide authentication and authorization services and is based on cookie technology. The system is designed to be implemented in a heterogenous environment with independent and diverse service providers. The communication between the servers is done via Web services. Additionally, plug-ins are available for other protocols that allow for easy integration of existing authentication and authorization components. A prototype system is operational at the Schroff Stiftungslehrstuhl Information Services and Electronic Markets.

1 INTRODUCTION

As more and more e-businesses and service providers emerge on the internet and especially the World Wide Web (WWW) the number of accounts and passwords a user has to handle and to remember increases rapidly. Many providers require some type of identification and personal data to offer their services. Users either tend to re-use the same account name and password for many different providers or write their credentials down and even carry them with them to have them always available. Both methods increase the risk that a malicious person might get access to personal data and abuse them which might cause an important damage to the user as well as the service provider. In addition, in decentralized organisations multiple user accounts increase IT administration and service costs.

Single sign-on (SSO) systems offer the possibility to use only one account for a multitude of distributed services (Shirey, 2000). The user logs into one of the services and is then able to access resources of other service providers without having to re-authenticate. The set of services for which SSO can be provided is called the SSO domain. SSO systems mainly stem from two sources, namely distributed operating systems and telecommunication infrastructure. Many different systems like Kerberos (Steiner et al., 1988) or Radius (Metz, 1999) have been developed that offer single sign-on for a special kind of environment. SSO solutions for distributed Web applications are special because they have to be based on standard internet mechanisms in order to be deployable. A multitude of user clients (Web browsers) exists on the market and there is no control over the choice of whichbrowser a user decides to use.

Besides the most often stated field of e-commerce, universities are a very promising domain of setting up SSO environments (Murawski, 2000; Anchan and Pegah, 2003). Nowadays extensive user profiles of students, teachers, and researchers are kept digitally and e-learning platforms and administrative systems require digital transmission of data. Universities not offering their students an online time-table system and access to lecture materials are even considered not to be up-to-date. Universities still represent a very heterogeneous system where each institute has already made its proper choice of platform and applications. It is difficult to impose common standards and protocols on these independent entities. Thus, to successfully implement an SSO system in such an

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 167–174.

environment it must be able to integrate different types existing systems and rely on standard mechanisms that can easily be adopted by different kinds of proprietary platforms.

We propose a single sign-on system for cross-domain authentication based on standard internet mechanisms that is designed in order to integrate various types of existing user accounts and user databases of service providers. The strength of our system lies in the clear separation of user credentials and user profiles from role assignment and access permission information. The communication between the different components of our system is based on standardized protocols to which existing applications can easily be adapted. The choice of policies is left to the service providers who can choose their level of trust towards other system components.

In the following sections, we first present our system and describe its functionalities. After that we discuss the characteristics of our system and compare it to other related systems for Web authentication.

2 PRESENTATION OF THE SYSTEM

The components and their communication protocols are shown in fig. 1. The system consists of one or multiple application servers that contain the Web applications a user wants to access. The application servers can reside in different domains. One or multiple authentication servers handle the authentication of the user. They can access one or more user databases where one database contains the data for one specific user domain. The authentication server is allowed to query only the credentials needed to authenticate the user (typically username and password, link (1) in fig. 1). The user database may contain additional personal data about the user which are only accessible by the application servers in the corresponding domain (link (3)). For each user database domain there is an authorization server that contains the role and access information for each user. Two kinds of roles are available: public roles are valid across domain borders and can be accessed by the authentication server (link (2)). Local roles are restricted to the respective domain and may be queried only by application servers of the corresponding domain (link (4)).

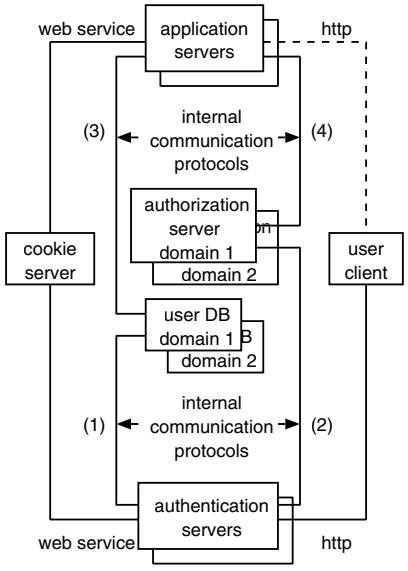


Figure 1: Components of the SSO system and communication links, dashed link: not necessarily secure.

The cookie server is the last element of the system. It contains information about the currently valid cookies, the corresponding users, and their public roles. The authentication servers may insert and delete entries in the cookie server database whereas application servers only perform entry look-ups. The description of the role system is not within the scope of this paper.

The communication between the user client and the application servers depends on the application requirements, it may or may not be secured. The remaining communication channels have to be secured as they carry sensitive data. The communication between the servers is mainly realized by the means of Web services over an SSL (secure socket layer) connection. Web services ease cross-domain access as they only require a standard HTTP connection. The communication with the user databases and the authorization servers may be adapted to the specific type of the underlying system (e.g. kerberos, secured Web service). Different plug-ins in the authentication servers offer customized communication. This allows for integration of Web applications with already existing authentication and authorization schemes.

2.1 Single Sign-on Procedure

Now how does the system offer single sign-on for multiple cross-domain applications? The procedure is the following (see also fig. 2, the letters refer to the messages in the figure):

1. The user is not logged on to any application. He sends a request for a service to application server *app1* (a).
2. *app1* cannot identify the user as no application server cookie is sent with the request. It there-

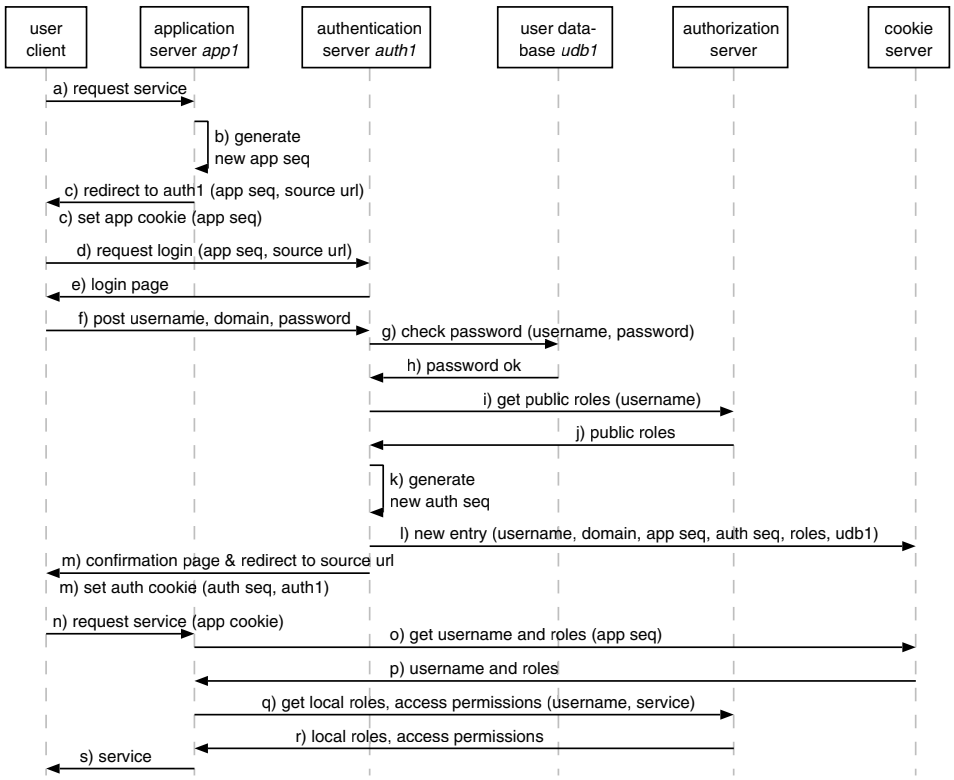


Figure 2: Sequence diagramm of SSO procedure at first login.

fore sends a redirection to one of the authentication servers, say authentication server *auth1*. The redirection contains the URL of the originally requested application page and a randomly generated sequence of characters, the application sequence. The application sequence is also included in the application cookie that is sent back to the user along with the redirect request (b, c).

3. *auth1* requests the credentials from the user consisting of a username, a password, and a domain (d, e). The credentials are transmitted over a secure connection (f).
4. *auth1* contacts the user database corresponding to the given domain to verify the correctness of the credentials (g). Once the verification succeeds (h) *auth1* retrieves the public role information for this user from the domain's authorization server. The roles are matched by means of the user name (i, j).
5. *auth1* randomly generates another sequence of characters, the authentication sequence (k). The sequences have to be long enough such that the probability of a malicious user reproducing a valid sequence by random trial is sufficiently small. The two sequences, the user name and domain, the list of public roles, and the name of the user database are transmitted to the cookie server who saves the information in its database (l).
6. An authentication cookie containing the authentication sequence and the identification of *auth1* is created and transmitted back to the user along with a login confirmation page. The page displays for some seconds and then redirects back to the originating application URL (m).
7. The user now has two cookies set: the application cookie and the authentication cookie. Upon the following request to *app1* the application cookie is transmitted (n). *app1* extracts the application sequence from the cookie and requests the user information from the cookie server (o).

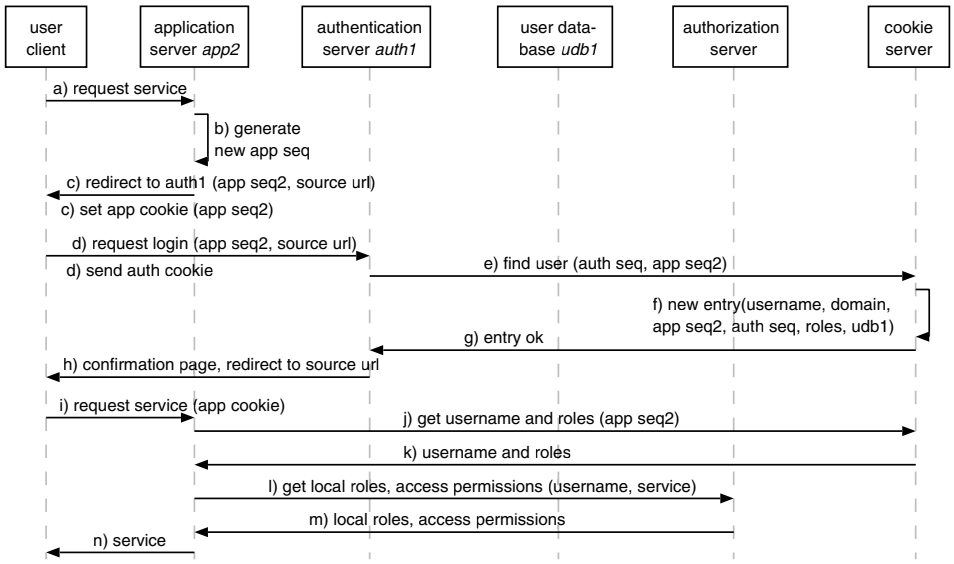


Figure 3: Sequence diagramm of SSO procedure at cross-domain service request.

8. The cookie server finds the existing entry containing the application sequence in its database and transmits the user name and his public roles to *app1* (p). *app1* may now look up local roles and access rights in the domain's authorization server and may deliver the requested service (q – s).

The user can now easily access services from all application servers in the domain of *app1* as the corresponding application cookie is always included in the requests. If a user now requests a service from a server residing in a different domain (say *app2*) the application cookie is not sent along with the request and the user is not recognized as logged in. The following steps are then performed (see fig. 3 and corresponding message numbering):

1. The user accesses *app2* without an application cookie (a). The server generates an application sequence (b) and redirects the user to *auth1*, setting the second application cookie (c).
2. *auth1* extracts the authentication sequence from the transmitted cookie (d) and requests the associated user information from the cookie server. The request also includes the new application sequence (e).
3. The cookie server finds the user associated with the authentication sequence and adds another entry in its database differing from the first only in the new application sequence (f, g).
4. *auth1* redirects the user who has now three cookies set back to the originating URL of *app2* (h). *app2* can now identify the user as described in the last two steps of the previous procedure (i – n).

In this scenario *app2* uses the same authentication server (*auth1*) as *app1*. But the system allows for several authentication servers in one SSO domain and each application server may choose which authentication server to use. For different authentication servers to recognize the user as logged in it is necessary that they be able to have a valid authentication cookie transmitted. As cookies (Kristol and Montulli, 1997) may not be set on behalf of other servers the following work-around is deployed. Each authentication server maintains a list of the other known authentication servers. After the first successful login the confirmation page displayed to the user contains for each other authentication server an HTML image tag including the authentication sequence and the name of *auth1*. The requests sent to the other servers allow them to set an own authentication cookie with the same authentication sequence. In reply to each request each authentication server delivers a transparent pixel image. Now the user can be recognized as logged in by all other authentication servers, as well.

However, steps 7 and 8 may increase the network traffic considerably as a request is sent to the cookie server upon each service request to an application server by a user. In order to minimize the network traffic and to increase the performance of the system the application server may locally cache the association of the application sequence to the username and roles. This design decision needs to be carefully considered with respect to the global logout procedure and its implications are discussed in the following section.

2.2 Global Logout

When the user logs out of one of the applications with which he is logged in (and of which he consequently possesses an application cookie) the logout has to be propagated to all other applications in the SSO domain, as well. Upon logging out the user is redirected to the authentication server. The authentication server sends a request to the cookie server to delete all entries containing the authentication sequence included in the authentication cookie. The authentication server sends a logout confirmation page to the user and deletes his authentication cookie. To delete the authentication cookies of all other authentication servers with which the user is currently registered the confirmation page contains again transparent images pointing to logout links of the other servers. Upon sending the images the other authentication cookies can be deleted, as well. If the user now accesses an application the application server will fail to find the corresponding entry in the cookie server and consequently can delete the application cookie. This is the standard case as described in the previous section.

If the application server uses caching to minimize network traffic the following two options of cache invalidation are possible:

1. The cache entries are only valid for a limited time and the application server checks the entries with the cookie server on cache expiration (variant 1).
2. The logout message is propagated to the application servers. This possibility is also employed in a similar manner in Microsoft's Passport SSO protocol (Microsoft Corporation, 2004b). It consists in keeping track of all applications the user is logged in with in the current session and to include transparent image requests for cookie deletion in the logout confirmation page, as well. The cookie server could easily keep track of the applications as a separate entry for each application sequence already exists. This would allow the application servers to locally cache user information and avoid subsequent requests to the cookie server while still validating the global logout without delay (variant 2).

2.3 Public Roles

Each user can be associated with several public and private roles. Public roles are valid across all applications in the SSO domain whereas private roles are local to an application or domain. A user's public roles are contained in the cookie server entries and can be read by all applications the user accesses. Each application may decide on its own whether to trust the validity of a public role.

Public roles can be used to identify users and user groups and their corresponding access rights across different applications. For example, if a student worker logs in at the site of the department where he works he may be assigned the private role "student worker" and the public role "student at faculty of mathematics". If he subsequently accesses the Web page of the faculty of mathematics he is recognized as a valid student by means of the public role and may be granted access to lecture material without the need to re-authenticate.

Public roles could also be used to identify single users across different domains without the need to exchange user names or to issue globally unique user identifiers. Each user may be assigned a public role corresponding e.g. to his matriculation number. With this information each student can be identified by all university applications whose user databases contain the matriculation numbers of their users.

2.4 Failure of Servers

For a single sign-on system it is important to note which of the participating servers constitute single points of failure. Unavailablilty of a server that provides necessary information to perform login functionality means that all services that require login and access control become unusable. It is therefore desirable to have different servers that are able to provide the same functionality. A slow authentication system due to increased server load is still preferable to a completely non-functional system.

In the system described in this paper the authentication server does not constitute a single point of failure. Many different authentication servers may be set up and as long as one of the authentication servers that an application server knows is functional the login can be performed. If an authentication server is temporarily not available this may result in a broken image tag displayed on the login confirmation page after a timeout but the login is still completed.

If a user database or an authorization server breaks down applications in the corresponding domain cannot be accessed but the other domains are not affected. In the current implementation only the cookie server as the only central component constitutes a single point of failure. By replication of the cookie

server's database and network interfaces this risk can be avoided, respectively reduced to a minimum. Mirror cookie servers do not increase the vulnerability of the overall system as the cookie server does not contain extremely sensitive data (like credentials, credit card information etc.). The sequences contained in the cookie server are only valid for a single session. To avoid replay of valid sequences in case that a malicious person gets access to the contents of a cookie server these could be stored encrypted either in the cookies or in the cookie server.

2.5 Scalability

For a single sign-on system to be practically relevant scalability is an important factor. An SSO system must be able to handle a large number of users, accounts, and applications without being considerably slowed down.

Looking at our solution, what does a large number of accounts and users imply?

- Accounts and their authorization information have to be stored in the user databases and authorization servers. As the system is already designed to have separate servers for each domain application specific scalability is not affected by the SSO procedure.
- The bottleneck of the current application is the central cookie server. The size complexity of the database is of the order *number of currently logged in users · number of currently used applications*. The number of concurrent users usually is only a small fraction of the total users, the same is true for the applications so that currently available data base systems should be able to accomodate cookie data both with respect to size and response time. The requests to the cookie server are all of the same structure such that efficient indexing techniques can be deployed. Tests with a single LDAP server containing about 25 million entries showed no performance degradation.
- Another important point is the amount of network traffic to and from the cookie server. For each login to a new application two requests to the cookie server are necessary (see fig. 2 and 3). The global logout requires one request for each authentication server the user logged in with. However, the most important amount of requests is generated at each access to an application service as the application server verifies the validity of the transmitted application cookie. These requests can be reduced or even avoided using the caching schemes as described in sec. 2.2 such that only a manageable amount of network traffic is left. Currently, the variant 1 described in sec. 2.2 is implemented.

2.6 The Role of Cookies for SSO

Building essential infrastructures on cookies has some important drawbacks. Many internet users have privacy concerns and do not want to use cookies, they can simply disable the cookie mechanism locally within their internet browser. There are also some security issues as cookies can on the one hand be modified by the user and on the other hand they may possibly be read or replayed by a third party. For a discussion on cookie security see e.g. (Samar, 1999).

Nevertheless, any SSO solution requires some sort of state management and user identification that cannot be provided by the stateless HTTP. Several possibilities exist to maintain information across multiple HTTP requests. An often used method is to include session information in the URLs, either as part of the query string or as an integral part of the URL itself. In contrast to cookies this does not require any data to be stored in files on the user's computer. This solution works very well for session management in one domain but it poses some problems when applied to cross-domain services. Each web server in an SSO domain would have to take care of including application and authentication sequences only in the URLs of the corresponding server. Otherwise sequences would risk to be exposed to third party websites who would be able to replay this information. Second, in our scenario web servers that issue redirect requests would have to have knowledge of the sequences that correspond to the redirect target server. In fig. 3, consider message (c). *app2* would have to include the authentication sequence of *auth1* in the redirect request. This requires additional transfer of sequence information over the network which means additional exposition of sensitive information to possible attackers. Thus we can see that replacing cookies with URL encoded sequences introduces additional security threats.

Keeping track of session information by using the user's IP address generally poses problems because of the use of proxy servers and network address translations, as well as the use of public terminals by many different users. Identification of the user through HTTP authentication also does not work for cross-domain scenarios. Other solutions require the user to install additional, specialized software and are no longer based on standard Internet mechanisms.

In general, privacy and single sign-on services are contradictory requirements that are difficult to combine. To minimize scepticism it is important to provide the user detailed information about what is stored in the cookies and what they are used for. Furthermore, an SSO solution based on cookies can still offer privacy concerned users alternative login mechanisms without cookies if they are willing to renounce single sign-on and log on to every service separately. In the current version, however, this is not yet implemented.

3 RELATED SYSTEMS

Many different solutions have been proposed in recent years for offering single sign-on in different environments (de Laat et al., 2000; Volchkov, 2001). Many of them are not transferable to a heterogeneous and distributed Web environment as they require a centralized structure of the involved components and common protocols. Other systems work with clients that are required to have additional functionality (Pfitzmann and Waidner, 2003) which we do not consider to be a realizable approach in the World Wide Web environment. In the following, we give an overview over the three solutions known to us regarding single sign-on for Web applications and compare the systems with our solution.

Microsoft's .NET Passport system (Microsoft Corporation, 2004b; Kormann and Rubin, 2000) is the largest functional single sign-on solution with 200 million accounts performing approximately 1350 authentication requests per second (March 13, 2003, (Microsoft Corporation, 2004a)). The system consists of a central Passport server that contains and manages all user accounts and corresponding user information and performs the authentication. Each user has a unique identifier. If a user has logged on to Passport from a service provider's site he can subsequently visit sites of other businesses adhering to Passport and is recognized by his unique identifier. The Passport server as well as the service providers set transient (session-only) or persistent cookies to recognize the user as already logged in. The cookies contain credential information about the user. A global logout from all Passport accounts is realized by including logout requests to all services the user is currently logged in with (and of which he possesses valid cookies) in HTML image tags.

The main difference between Passport and our solution is Passport's central authentication and user database server. This server is a single point of failure, if the service breaks down all businesses using Passport are unavailable to the users. It is not stated which measures are taken to offer scalability and backup servers. Possible dangers of replicated databases are discussed in (Kormann and Rubin, 2000). Our solution works with different, distributed authentication servers and integrates multiple, local user databases without requiring a separate, unique identifier. Identification across multiple domains may be realized at different levels by the means of public roles. No credential information is stored in local user cookies which may potentially be modified by the user himself and is not well protected against access from intruders.

The Liberty Alliance Project (Liberty Alliance Project, 2003) was formed in September 2001 by an association of several major companies in order to specify open standards for federated network identity management. The architecture (Liberty Alliance Project, 2003) allows for each service provider to maintain its own user database and user identities. A user who wants to use single sign-on between service providers with different identities may select to federate these identities between the service providers who are now able to match the foreign identity to their own. Without federation the user has to separately log into each service provider. The decision of whether to trust a user logged in with a different, federated identity remains the service provider's local policy. Identity providers take the role of the authentication server and the cookie server in our scenario, they use cookies to maintain a user's login state.

Like our solution, the Liberty architecture allows for integration of existing user databases and accounts of different service providers. The choice of whether to trust a federated identity can be made locally by each service provider, like it is the case with the public roles introduced in our system. The server architecture is nevertheless different. Liberty does not state whether SSO is possible across multiple identity providers and does not distinguish between authentication server and cookie server. Having the possibility to choose a nearby authentication server reduces the route length over which passwords and other credential information have to be carried to a minimum whereas no such sensitive data has to be transferred to and from the cookie server. Data have to be transmitted between the cookie and the authentication server only upon an initial service provider login. It is therefore useful and improves the security of the system to introduce this flexibility without having an important impact on the overall system performance.

Samar (Samar, 1999) proposes different protocols for cookie-based single sign-on. For cross-domain authentication, he introduces two centralized servers, a cookie server and a login server. The login server contains the authentication information about all users in the SSO domain and recognizes signed-in users by means of a login server cookie. The cookie server contains information about the different Web application cookies. Samar does not give detailed information about the possibility of a global logout in this scenario.

Samar's approach does not offer the integration of decentralized user databases and introduces two single points of failure, namely the login and the cookie server. The approach is similar to Microsoft's Passport except that Passport integrates the two different

server types in one central server. Authorization is done locally at the service providers.

4 CONCLUSION

In this paper we propose a single sign-on system architecture based on standard HTTP mechanisms for distributed, cross-domain Web applications. The system allows for integration of distributed, proprietary user databases and authorization servers. User accounts need neither to be centralized nor to be explicitly exposed to other service providers in order to provide SSO services. Public roles offer a flexible mechanism to transport user information across different domains while leaving the final decision of whether to accept public roles to the single service providers. The system works with multiple, cross-domain authentication servers and a centralized cookie server. Different possibilities for implementing a global logout are proposed. We compared the system to other, well-known propositions and solutions for offering cross-domain single sign-on in a world wide Web environment and discussed similarities and differences between the systems as well as advantages and drawbacks.

The prototype system is already functional over different domains at the department *Information Services and Electronic Markets*. To test the system please visit http://demo.em.uni-karlsruhe.de/sso/ and http://sso.itloesungen.com/, log into one of them as indicated on the login screen, and test the system by visiting the other. First experiences with the system revealed no major difficulties concerning usability and the initial user reaction towards the single sign-on service was very positive.

The next step is to integrate other domains of university institutes and departments, affiliated companies and student organisations to test the system at a larger scale and to identify possible improvements especially concerning the adaption of proprietary systems into the SSO environment. Further research will be directed towards the analysis of various attack scenarios and on role contracting models.

REFERENCES

Anchan, D. and Pegah, M. (2003). Regaining single sign-on taming the beast. In *Proceedings of the 31st annual ACM SIGUCCS conference on user services*, pages 166–171.

de Laat, C., Gross, G., Gommans, L., Vollbrecht, J., and Spence, D. (2000). RFC 2903: Generic AAA Architecture. Network Working Group.

Kormann, D. P. and Rubin, A. D. (2000). Risks of the passport single signon protocol. *Computer Networks*, 33:51–58.

Kristol, D. and Montulli, L. (1997). HTTP State Management Mechanism. Network Working Group RFC 2109.

Liberty Alliance Project (2003). Liberty Architecture Overview v1.1. Technical report, Liberty Alliance Project. http://www.projectliberty.org.

Metz, C. (1999). AAA protocols: Authentication, authorization and accounting for the internet. *IEEE Internet Computing*, 3(6):75–79.

Microsoft Corporation (2004a). Microsoft .NET Passport for Businesses. http://www.microsoft.com/-net/passport/services/business.asp, accessed Feb 25, 2004.

Microsoft Corporation (2004b). .NET Passport Review Guide. Technical report. http://www.microsoft.com/.

Murawski, R. (2000). Centralized directory services and accounts management project. In *Proceedings of the 28th annual ACM SIGUCCS conference on User services: Building the future*, pages 198–201.

Pfitzmann, B. and Waidner, M. (2003). Analysis of liberty single sign-on with enabled clients. *IEEE Internet Computing*, 7(6):38–44.

Samar, V. (1999). Single sign-on using cookies for web applications. In *Enabling Technologies: Infrastructure for Collaborative Enterprises*, pages 158–163. IEEE.

Shirey, R. (2000). Internet security glossary. Network Working Group RFC 2828.

Steiner, J. G., Neumann, B. C., and Schiller, J. I. (1988). Kerberos: An authentication service for open network systems. In *Usenix Conference Proceedings*, pages 191–202.

Volchkov, A. (2001). Revisiting single sign-on: A pragmatic approach in a new context. *IT Professional*, 3(1):39–45.

PART 3

Wireless Communication Systems and Networks

ADJACENT CHANNEL INTERFERENCE
Impact on the Capacity of WCDMA/FDD Networks

Daniel Figueiredo, Pedro Matos, Nuno Cota and António Rodrigues
Instituto Superior Técnico, Technical University of Lisbon, Av. Rovisco Pais, Lisbon, Portugal
Email: danif@netcabo.pt, pvgmatos@hotmail.com, ncota@isel.ipl.pt, antonio.rodrigues@lx.it.pt

Keywords: WCDMA, adjacent channel interference (ACI), FDD, spectrum management.

Abstract: The adjacent channel interference (ACI) can result in a reduced network capacity in a multioperator WCDMA/FDD environment. This paper is devoted to the study of the ACI, using a static simulator. Simulations were performed in order to identify particular scenarios and network compositions where ACI plays a major role in the system capacity. On the basis of the results, the authors identify the best strategy for frequency deployment within the available spectrum. It is demonstrated that the macro carrier should be located in the centre of the frequency band, protected from the ACI introduced by other operators. It is, in fact, the carrier which suffers the greatest losses caused by the increase in ACI. Furthermore, the micro carrier should be placed as close as possible to the adjacent channel of other operators in order to maximize system capacity.

1 INTRODUCTION

At the moment radio spectrum is becoming increasingly more occupied, making its management a vital tool to network planning. It is under these circumstances that the third generation (3G) mobile communications systems emerged, and in particular the UMTS (Universal Mobile Telecommunications System) in Europe. The air interface chosen for the 3G UMTS system was WCDMA (Wideband Code Division Multiple Access) for the paired bands FDD (Frequency Division Duplex). The scope of this paper is to study the impact of frequency utilization on WCDMA/FDD networks and develop a strategy to optimise it.

The performance of any CDMA system is conditioned by the interference. From the various possible sources of interference that are present in these systems this paper focuses on the study of adjacent channel interference (ACI) and its effect on the overall capacity of the system.

This study consists of five main sections. Following a brief overview of a few features of UMTS related with the options chosen for the simulations, to be found in Section 2, the criteria for the choice of the tested scenarios are explained in Section 3. Section 4 presents the results obtained with the simulated scenarios, and final conclusions are drawn in Section 5.

2 INTERFERENCE ISSUES IN UMTS FDD NETWORKS

When defining the UMTS system, 3GPP (3rd Generation Partnership Project) inferred that each radio channel has a bandwidth of 5 MHz and the channels allocated are positioned beside each other in uplink and downlink bands separated by 190 MHz (3GPP 25.101). Each channel has its carrier and these are assigned to the UMTS operators in the market.

The strategy designed to reduce the interference, thus achieving the highest possible capacity, consists in identifying optimal spacing between carriers in the radio spectrum available for use. For this purpose, a specific frequency arrangement must be considered, as it may vary from country to country.

The initial licensed spectrum for UMTS in FDD mode was a band with twelve carriers, both uplink and downlink. The case considered includes four operators, each of which has been allocated three carriers. Within the allocated band, it is possible to

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 177–183.

choose the spacing between its carriers and the distance from adjacent operator carriers. 3GPP defines for the spacing between channels a raster of 200 kHz (3GPP 25.104), which means that the spacing between carriers can vary in increments of 200 kHz around 5 MHz.

As seen in Figure 1, in order to decide which spacing should be used, more issues must be taken into account, to prevent the carriers from encroaching on their neighbours. Consequently, bearing in mind the limits typically used in simulations, it has been chosen to vary the distances between the values 4.6 and 5.2 MHz.

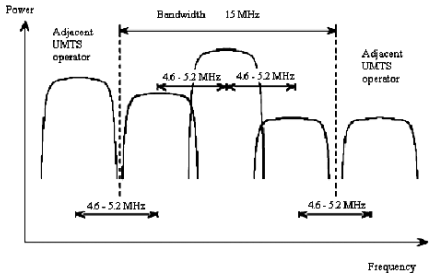


Figure 1: Spacing between carriers in a UMTS system (adapted from (Holma and Toskala, 2002)).

In a WCDMA system, developed in the above context, the interference can stem from a large number of sources, namely, thermal noise, traffic in the same cell, traffic in adjacent cells and traffic from operators using adjacent cells.

Possible ways of measuring the interference leakage between connections operating on different carriers must be considered. As the filter is not perfect, when transmitting in its own channel, one carrier will send part of its power into adjacent channels. This effect is measured as the ACLR (Adjacent Channel Leakage Ratio). On the other hand, the receiver filter is unable to receive only the desired signal alone, which is why the rejection of the adjacent channel signal is measured as ACS (Adjacent Channel Selectivity). Moreover, when considering the existence of two carriers which interfering with each other, the total interference is given as an ACIR (Adjacent Channel Interference Ratio) and determined by (1).

$$ACIR = \frac{1}{\frac{1}{ACLR} + \frac{1}{ACS}} \quad (1)$$

Furthermore, this source of interference can be seen both from the uplink and from the downlink standpoint. Consider an uplink connection, whose ACIR is given in (2) below. As it is quite likely that the filter in the user equipment (UE) will be poorer than the filter in the base station (BS), the UE ACLR dominates in the case of uplink. In downlink, the situation is analogous, as seen in (3), where the UE ACS dominating on this occasion.

$$ACIR_{UL} = \frac{1}{\frac{1}{ACLR_{UE}} + \frac{1}{ACS_{BS}}} \quad (2)$$

$$ACIR_{DL} = \frac{1}{\frac{1}{ACLR_{BS}} + \frac{1}{ACS_{UE}}} \quad (3)$$

One of the essential parameters used in the simulation was that the value of the filter depends on the spacing given to the channels. 3GPP defines the filter's mask, while identifying minimum values for filters at 5 and 10 MHz [3, 4]. However, in this project more realistic values were used, which correspond to real equipment presently available. These values may be found in Table 1.

Table 1: Values of the filters used in BS and UE.

Spacing (MHz)	ACLR (dB)		ACS (dB)	
	UE	BS	UE	BS
5	33	60	33	70
10	43	65	43	

When simulations were run using spacing different from 5 or 10 MHz, e.g. 4.6 MHz, a logarithmic regression is made to convert the filter value and obtain a valid method to compare the results.

3 SIMULATION SCENARIOS

The choice of which scenarios to study was not as simple as it might be assumed at first glance. One of the goals in this paper was to find scenarios where ACI has a major role on the network's capacity, in order to understand the impact of placing carriers with different spacing. In (3GPP 25.942), the authors give an idea of the issues to be born in mind when choosing which scenarios to simulate.

In a preliminary stage, the search started with the study of the representative scenarios of rural and urban environments. When simulating two operators, BSs working with adjacent carriers were uniformly distributed over a map. In order to

simulate a worst case situation, the sites of both operators are not co-located and the interoperator spatial offset is equal to the cell radius (Hiltunen, 2002).

It was found that the inter-frequency interference impact on the capacity was minimal, when compared with the intra-frequency interference. The reason for this result lies in the fact that there are too many BSs from the same carrier interfering with each other.

The next step taken was to identify scenarios where the ACI played a significant role, at least as important as the interference coming from the connections working on the same carrier. Following simple scenarios, where just a few BSs and two carriers were taken into account, the analysis developed to encompass broader environments simulating urban centres with many antenna sites and three carriers.

A simple map was used as an entrance parameter to the simulator, with no additional information, apart from UE and BS positions. When placing the BSs of two different carriers, one must decide whether they are co-located, i.e. both cells lie on the same site, or not. In the latter situation, it is assumed that the worst case for ACI happens, i.e. the adjacent channel site is located at the coverage edge of the first channel cell.

The simulator used to achieve this analysis was static, using a Monte-Carlo evaluation method. As a result, the users were placed randomly on the map. Following the iterative process, only the connections with sufficient Eb/No (or signal to noise ratio - SNR) for the appointed service were considered to be served by the system. This simulator was adapted from the previous one described in (Laiho et al., 2002) and (Wacker et al., 2001). By examining many static situations, referred to as snapshots, network capacity is estimated through the average number of the served users (Povey et al., 2003).

The bit rates tested in this study were chosen in accordance with the services expected to be offered by operators in the first implementation phase. In this case, 12.2 kbps with CS (Circuit Switching), 64 kbps with CS and, finally, 128 kbps in downlink and 64 kbps in uplink using PS (Packet Switching). The results are presented taking into account users accessing one of these three types of services.

The UE power classes considered for determining the maximum output power were class 3 (24 dBm) for voice and class 4 (21 dBm) for data services (3GPP 25.101). The BS maximum output power used was 43 dBm.

Two types of antennas were chosen to simulate macro and micro BS: for the macro BS, tri-sectorized antennas with 18 dBi of gain, and for the micro BS, omni-directional antennas with 4 dBi of gain.

Two different propagation models were considered to calculate the path loss according to the characteristics of the environment (both for outdoor propagation). For rural scenarios the COST 231 Hata model was used. The main input parameters for the model are the UE antenna heights, 1.5 m, and BS antenna heights, 35 m. For the urban environment the propagation model applied was COST 231 Walfish-Ikegami. The main parameters used are UE antenna heights, 1.5 m, BS antenna heights, between 10 and 25 m (depending if they are macro or micro), street width, 20 m, building separation, 40 m, and building height, 12 m.

4 RESULTS

In the extended study that originated this paper, a wide range of scenarios and environments were considered (Figueiredo and Matos, 2003). Urban, rural and motorway environment were tested using layers containing twenty-three macro cells placed in the form of a grid. Furthermore, eight scenarios with only a few antennas (five at the most) were run to simulate specific situations using macro and micro cells. The dense urban environment was simulated, by using macro cells layers and micro cells to cover identified hotspots. In this paper, only the three most significant tests will be presented.

At the end of each simulation, the outputs were analysed. Apart from the maps indicating the BS and UE position, the network's capacity (measured in average number of served users) and the capacity loss (when compared with no ACI), parameters like the ratio between sources of interference were also analysed in (Figueiredo and Matos, 2003). The interference sources considered in the results included the interference coming from the adjacent channel, the interference from the same channel from neighbouring cells and the interference from the same cell (due to the other users connected to the same BS).

A maximum load of 50% was allowed in the radio interface.

4.1 Case 1: Small Scale Networks with Two Operators

This scenario was developed to study the impact of a new micro BS placed by an operator to cover a hotspot in the middle of an existing network of macro BSs from the adjacent carrier competitor.

The users from both operators have been located around the centre of the area considered. The area simulated has a high density of active clients, as shown in Figure 2.

The number of users presented in Figure 2 corresponds to the initial users from each operator, and are placed on the map at the simulation start.

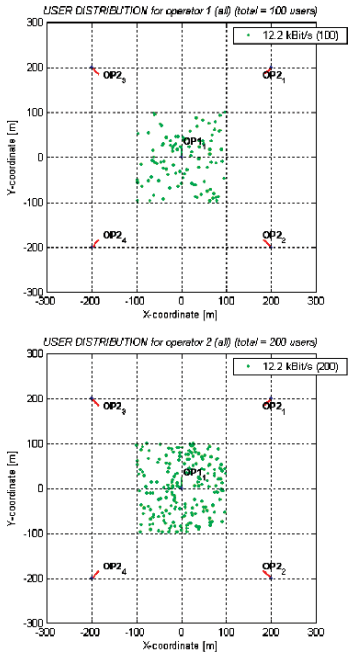


Figure 2: Location of UE (CS 12.2) and BS (Case 1).

Table 2 shows the average number of users served employing different channel spacing and the percentage of loss compared to the case where no adjacent operator exists (without ACI) for the CS 12.2 kbps service.

It has been verified that the operator 2, covering the area with four macro BSs, is the one that suffers most from interference. This may be explained by the fact that users of operator 1 (micro BS) are closer to the antenna, which therefore makes it more difficult for them to lose the connection. A comparison of the results obtained from the simulations performed with the three services tested for operator 2, is shown in the graph presented in Figure 3.

Table 2: Results from the simulated scenario (Case 1).

CS 12.2 kbps	Capacity (average number of users)	Capacity Loss (%)
Operator 1		
Without ACI	77	0
4.6 MHz	76.1	1.17
4.8 MHz	75.8	1.56
5.0 MHz	76.7	0.39
5.2 MHz	76.7	0.39
10.0 MHz	77	0
Operator 2		
Without ACI	162.1	0
4.6 MHz	44.8	72.36
4.8 MHz	104.7	35.41
5.0 MHz	126.2	22.15
5.2 MHz	134.6	16.96
10.0 MHz	155.5	4.07

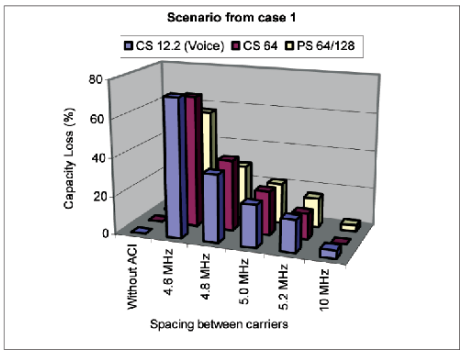


Figure 3: Capacity Loss of operator 2 (Case 1).

4.2 Case 2: One Layer Macro and Two Layers Micro

In the situation shown in Figure 5 several macro BSs were placed to form a grid and cover the area to serve users of operator 1. Four hotspot areas (with higher user density) from both operators 1 and 2 were placed and micro BSs located to cover them.

In this case, it has been tested an environment where three carriers coexist and interfere with each other. Operator 1 has one carrier for macro BSs (f1) and another for micro BSs (f2). Operator 2 has only one carrier for micro BSs (f3). The three channels were allocated next to each other in the radio spectrum as shown in Table 3.

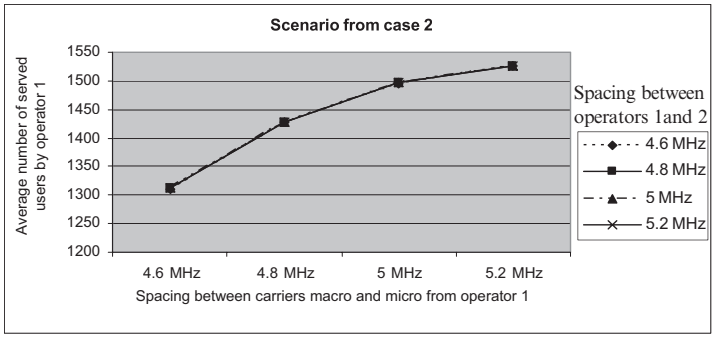


Figure 4: Capacity of operator 1 (Case 2).

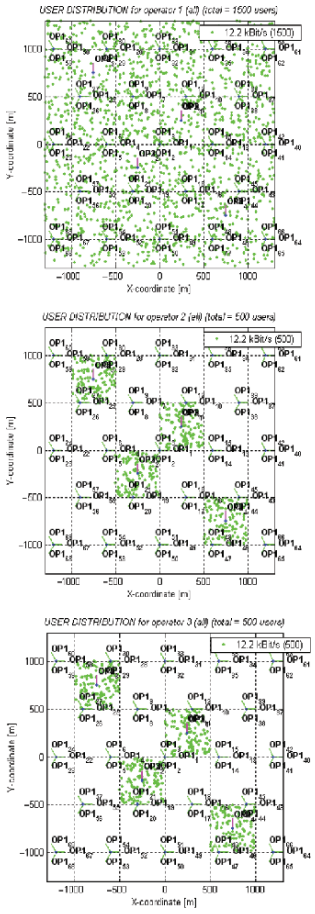


Figure 5: Location of UE (CS 12.2) and BS (Case 2).

Table 3: Frequency planning with macro (M) and micro (m) networks (Case 2).

Operator 1		Operator 2
f1 (M)	f2 (m)	f3 (m)

Once again the spacing considered between each of the three channels varied within the range 4.6 to 5.2 MHz.

The results obtained from these simulations are given in Figure 4. The graph shows the average number of users served by operator 1, and take into account both the users connected to the macro (f1) and to the micro (f2) layers. As expected, it can be seen that the network capacity rises when the spacing between f1 and f2 widens. As both micro layers accommodate fewer users than the macro layer from operator 1, it is evident that the distance between micro layers (f2 and f3) from the different operators does not have a great impact on the capacity of operator 1. This fact is confirmed by the graph, since the lines are almost overlapped.

Following the analysis of these results, it is logically preferable to choose a wider spacing between carriers f1 and f2, in order to achieve an increase in the capacity of operator 1. At the same time, it is reasonable to leave the lowest distance to the carrier from operator 2 (f3), since the damage is imperceptible, towards optimisation of the spectrum allocated.

4.3 Case 3:Two Layers Macro and One Layer Micro

As in the previous case, in Case 3, the authors tested the impact of interference among three carriers controlled by two operators. However, in this case, there are two major macro BSs grids from operators 1 and 2. The same hotspots mentioned in

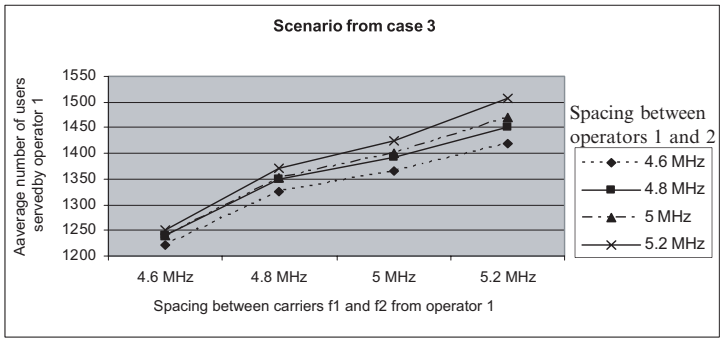


Figure 6: Capacity of operator 1 (Case 3).

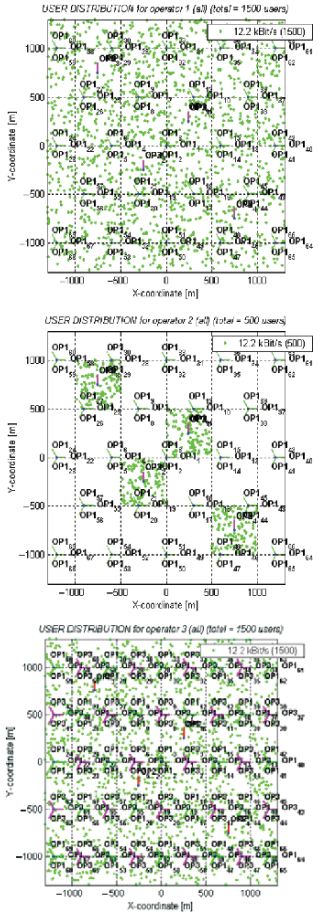


Figure 7: Location of UE (CS 12.2) and BS (Case 3).

Case 2 are now covered with micro BSs by operator 1 only, as seen in Figure 7. Thus, the configuration of the radio spectrum is similar, the only difference being that there are two channels for macro BSs and one for micro BSs, as shown in Table 4.

Table 4: Frequency planning with macro (M) and micro (m) networks (Case 3).

Operator 1		Operator 2
f1 (M)	f2 (m)	f3 (M)

The results, presented in the graph from Figure 6, were obtained by using the same procedure followed in Case 2.

As before, the network's capacity grows as the distance between f1 and f2 becomes larger. However, it can now be confirmed that the spacing between the two adjacent carriers from different operators (f2 and f3) has a significant impact on the overall capacity of operator 1. This feature is due to the fact that carrier f3, from operator 2, now accommodates a much larger number of users in its macro layer.

In this situation the analysis has to be considered more carefully than in the previous case. To achieve maximum capacity in operator 1, apparently the best solution would be to choose the maximum spacing between carriers f1 and f2 whilst, at the same time, also leaving the highest distance to the adjacent operator channel (f3). However, in doing so, one is failing to take into account the fact that each operator has three allocated channels. Note that, in the future, it will be valuable to use all of them to face an anticipated traffic increase. As a result, it would wise to choose a configuration in which the two carriers from operator 1 are not positioned in such a way that they occupy the free space left by hitherto unused third channel.

5 CONCLUSIONS

In this paper the authors studied the impact of the ACI on a general network's capacity. This led to some more useful conclusions that may be applied when planning the launch of a WCDMA/FDD radio networks.

When considering two wide BSs grids that lie close to each other to cover a specific area, it was observed that the main interference source is not the ACI, but interference from the neighbouring BSs, working on the same channel.

From scenarios like the one presented in Case 1, it was noted that the macro BSs are more likely to suffer from ACI when new hotspots are covered with micro BS by a competitor operator. This fact is explained by the longer distance between the user and the macro BS, as compared with the latter. As the macro carrier may suffer a greater impact on capacity, it should be protected and placed in the centre channel of the allocated spectrum. This choice is irrespective of the number or type of carriers used, assuming that the operator launching a service uses at least one macro carrier.

It may also be seen, from these small and specific case scenarios like Case 1, that the use of a 4.6 MHz spacing may cause critical problems, leading to a serious reduction in the network's capacity. Therefore, distances between carriers of 4.6 MHz or less should never be used. Although in the vast majority of the situations the loss may not be that disastrous, the possibility of having certain areas with losses above 50% is unsustainable to an operator.

Upon considering an available spectrum of three carriers, and assuming that the macro carrier is located in the centre channel, it is intended to decide where to place the micro channel. From Case 2, where operator 2 placed a micro carrier in the channel adjacent to the spectrum of operator 1, it was seen that the distance between the two channels was almost irrelevant to the overall network's capacity. However, when operator 2 has a macro carrier on the channel, adjacent to operator 1, the latter suffers the consequences of a reduction in the distance between different operators' channels. On the basis of the compromise solution of not occupying the spectrum of the three channels allocated using two carriers only, it may be concluded that a spacing of 5.2 MHz between f1 and f2 and 4.8 MHz between f2 and f3 is the best option.

ACKNOWLEDGEMENTS

We would like to thank Luís Santo and Ana Claro for their support and useful discussions that helped to improve this work. The authors are also grateful to *Optimus* for the support given to this project.

REFERENCES

Laiho, J., Wacker, A., and Novosad, T., 2002, *Radio Network Planning and Optimisation for UMTS*, John Wiley & Sons, Sussex, England

Holma, H., and Toskala, A., 2002, *WCDMA for UMTS – 2nd Edition*, John Wiley & Sons, Sussex, England

3GPP Technical Specification 25.101 v5.5.0, UE Radio Transmission and Reception (FDD)

3GPP Technical Specification 25.104 v5.5.0, BS Radio Transmission and Reception (FDD)

3GPP Technical Specification 25.942 v5.1.0, Radio Frequency (RF) System Scenarios

Hiltunen, K., 2002, *Interference in WCDMA Multi-Operator Environments*, Postgraduate Course in Radio Communications 2002-2003, Helsinki University of Technology, Finland

Wacker, A., Laiho, J., Sipilä, K., Heiska, K., and Heikkinen, 2001, K., *NPSW – MatLab Implementation of a Static Radio Network Planning Tool for Wideband CDMA*

Povey, G., Gatzoulis, L., Stewart, L., and Band, I., 2003, *WCDMA Inter-operator Interference and "Dead Zones"*, Elektrobit (UK) Ltd, University of Edinburgh

Figueiredo, D., and Matos, P., 2003, *Analysis, Impact and Strategy of Frequency Utilisation on WCDMA/FDD Networks* (in Portuguese), Final Graduation Thesis, Instituto Superior Técnico, Lisboa, Portugal

CARE-OF-PREFIX ROUTING FOR MOVING NETWORKS IN MOBILE IP NETWORK

Toshihiro Suzuki, Ken Igarashi, Hiroshi Kawakami and Akira Miura
NTT DoCoMo, 3-5, Hikarinooka, Yokosuka-shi, Kanagawa, 239-8659 Japan
Email:{toshi, igarashi, kawakami, miura}@netlab.nttdocomo.co.jp

Keywords: Mobile IP, moving network, routing, addressing, handoff, NEMO.

Abstract: The future ubiquitous network will serve so many mobile terminals that it is extremely important to control them efficiently. One useful approach is to group terminals having similar movement characteristics and manage them in units of groups. Another important issue is the mobility management of moving networks, such as a network on a train or in a car, or a personal area network. Moving networks may be defined for a variety of situations and can lead to a lot of attractive applications. Moving network mobility support is indeed one of the most interesting research topics. In this paper, we clarify the difference between host mobility support and the conventional moving network mobility support, propose a mechanism for moving network mobility support and shows it is better than the conventional ones.

1 INTRODUCTION

Since the future ubiquitous network must serve several billion Mobile Nodes (MNs) (i.e., mobile terminals), it is extremely important to control them efficiently. Given this number of MNs, one key technique is to group MNs having similar movement characteristics, and manage them in units of groups. Another urgent topic is to enhance the mobility management of local networks, such as a network on a train, in a car, or a personal area network. This moving network mobility support and moving networks can be applied to a variety of situations and can lead to a lot of attractive applications. This mobility management is indeed one of the most interesting research topics today. Many groups including *IETF* are actively researching IP routing techniques to support moving network mobility. NTT DoCoMo Network Laboratories are also studying it as a key technology for IP^2 (IP based IMP Platform) (Yumiba, 2001), a platform we have proposed for the next-generation mobile network.

The representative requirements for moving network mobility support in IP are the same as those for host mobility support (mobility management for moving hosts rather than a moving network). They are:

(1) Route optimization
(2) Minimization of the packet header size
(3) Reduction in handoff signal overhead.

"Pinball" Routing (Thubert, 2004), in which packets are always transmitted via Home Agent (HA) (Johnson, 2004), cannot satisfy requirement (1) because it requires excessive network resources. Given requirement (2), we must minimize packet overhead by dispensing with encapsulation. Requirement (3) demands that handoff be achieved seamlessly with minimal packet loss and short handoff latency. Therefore, it is important to reduce the amount of handoff signals. The *NEMO WG* has proposed only partial solutions to these three requirements.

In this paper, we clarify the difference between host mobility support and the conventional moving network mobility support, and propose a solution that satisfies all the requirements. Its effectiveness was confirmed by using *network simulator 2* (called ns2).

Section 2 briefly describes the difference between host mobility support and the conventional moving network mobility support, and the requirements for moving network mobility support. Section 3 proposes the basic techniques of a new routing method applicable to moving networks. Section 4 introduces a new routing mechanism that uses these basic techniques for Mobile IP (MIP) (Johnson, 2004). Section 5 compares the proposed routing mechanisms with conventional ones.

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 185–191.

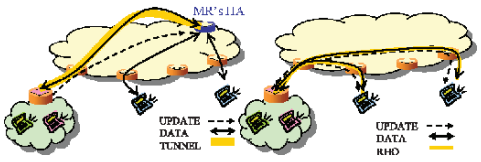


Figure 1: Basic. Figure 2: RRH.

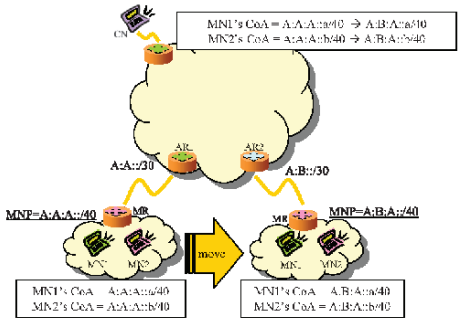


Figure 3: Care of Prefix.

2 DIFFERENCE BETWEEN HOST MOBILITY SUPPORT AND THE CONVENTIONAL MOVING NETWORK MOBILITY SUPPORT

The characteristics and brief evaluations of NEMO Basic Support (hereafter referred to as Basic) (Devarapalli, 2004) and Reverse Routing Header (RRH) (Thubert, 2004), both of which are currently proposed in NEMO WG for moving network mobility support, are shown below.

Basic constructs a bidirectional tunnel between a Mobile Router (MR) and the HA of that MR. Packets from/to MNs in a moving network are always carried via this tunnel (Fig. 1). When the moving network moves, handoff is achieved by reconstructing a tunnel. Specifically, a bidirectional tunnel is reconstructed between the Care of Address (CoA) (Johnson, 2004), which an MR is allocated by the new AR (Access Router), and the HA address of the MR. CoAs of MNs in the moving network remain unchanged even if handoff occurs. This hides the move of the moving network from the MNs in the moving network. Furthermore, even if there are many MNs in the moving network, handoff can be achieved easily because only this bidirectional tunnel needs to be reconstructed. Therefore, there is a high possibility that requirement (3) can be met. However, the undesirable effect of Pinball Routing is significant if the HA of the MR is far from the moving network. Additionally, packet overhead is greatly increased because packets are doubly encapsulated, once for the bidirectional tunnel and another for the CoA of the MN. Therefore, Basic cannot meet requirements (1) and (2).

RRH satisfies requirement (1) (Fig. 2). Specifically, routing is optimized as follows. All CNs are informed of the CoA of the MR, and the packets destined to MNs in the moving network are transmitted with Routing Header Option (RHO) in IPv6 (Deering, 2004). That is, the CoA of the MR and the CoA of the MN are attached. With regard to Requirement (2), RRH yields packet header sizes that lie between those of Basic and host mobility support. When the moving network moves, it is necessary to inform all CNs of the change in the CoA of the MR. The more CNs there are, the more handoff signals are sent. Therefore, RRH cannot meet requirement (3).

As mentioned above, neither of the two conventional mechanisms can satisfy all requirements. This is due to the assumption made by NEMO WG for the moving network. NEMO WG assumes that the prefix inside a moving network, i.e. Moving Network Prefix (MNP), is fixed (Devarapalli, 2004) (Thubert, 2004). Given this assumption, MNP is not changed even if a moving network moves. Therefore, to connect to an MN in the moving network, it is necessary to use the CoA of the MR in addition to the CoA of the MN in the moving network. This increases the packet header size. The MR-CoA is needed to construct a bidirectional tunnel in Basic, or to set it in the RHO in RRH.

The technologies proposed in Section 3 dispense with this assumption. That is, MNP is changed to adapt to the hierarchical address for the AR to which the moving network is connected. This enables packets to be routed to an MN in the moving network using only MN-CoA in the moving network, as in the case of host mobility support.

3 PROPOSED BASIC TECHNIQUES

3.1 Care of Prefix

As described in Section 2, the conventional mechanisms require the use of both of MR-CoA and

MN-CoA to route packets to an MN in the moving network, which increases packet overhead. Therefore, it is necessary to find a solution that minimizes packet overhead. The solution should be to use only one CoA, as in the case of host mobility support. The Care of Prefix (CoP) (Suzuki, 2003) technique is used to implement this. Specifically, an MR is allocated a CoP by the AR to which the moving network is connected. This CoP is an MNP in the hierarchical topology that embraces the moving network. After that, the MR uses this CoP to

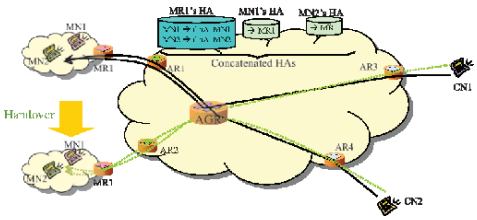


Figure 4: Concatenated HAs and Aggregate Router.

assign a CoA to the MNs in the moving network. In this way, packets for an MN in the moving network can reach the MN using only MN-CoA (Fig. 3).

The method of allocating the CoP is shown in Figure 3, using the IPv6 method as an example. Suppose that the net mask of the AR, which is an edge router of the core network, is 30 bits long. The moving network is allocated a CoP with a 40-bit mask to form a hierarchical structure that embraces the moving network.

In this way, MNP (i.e., CoP) reflects the hierarchical topology of the core network so that MN-CoA can be resolved from anywhere within the core network. In addition, a CoA can be generated from a CoP without any risk of duplication. Since the CoP is uniquely allocated to each moving network, duplicate CoAs are not generated for MNs that are connected to the same AR.

CoP makes it possible to meet requirements (1) and (2) at the same time because a CN can directly send packets to an MN in a moving network using only MN-CoA in the same manner as in host mobility support. However, when handoff occurs, the CoAs of all MNs in a moving network must be changed. This dramatically increases the number of handoff signals sent to the HAs of all MNs, and similarly the number of those sent to all CNs if route optimization is implemented. Therefore, it is difficult to meet requirement (3).

3.2 Concatenated HAs

As mentioned in Section 3.1, the use of CoP cannot meet requirement (3). One problem is that handoff signals must be sent to the HAs of all MNs in a moving network. To solve this problem, we propose Concatenated HAs (Suzuki, 2003) (Fig. 4).

In this technique, each HA of each MN does not hold its CoA. Instead, it holds the information that the MN is in a certain moving network. Specifically, the information of MN-MR concatenation is registered with the HA of each MN, while the CoAs of all MNs are registered with the HA of that MR. This makes it possible to limit the number of entities updated at handoff. At handoff, only the HA of the MR requires updating rather than the HAs of all MNs.

3.3 Aggregate Router

As mentioned in Section 3.1, there is another problem that prevents requirement (3) from being satisfied. It is that handoff signals must be sent to all CNs. To solve this problem, we propose the AGgregate Router (AGR) (Fig. 4). The purposes of the AGR are twofold: localize handoff signals and aggregate the handoff signals that are sent to all CNs. Specifically, the AGR manages the mobility of the moving network as well as the HA of MR, i.e., the AGR maintains the CoAs of all MNs in the moving network, and each CN holds the binding information that indicates that MN-CoA is the AGR address. If the CoAs of all MNs in the moving network are changed due to handoff, the MNs do not need to send handoff signals to each CN. They only send handoff signals to the AGR. This localizes the handoff signals. Furthermore, we aggregate them if MR sends a handoff signal to AGR instead of all MNs. Moreover, the binding information that MN-CoA is AGR address can also be registered at each HA of each MN in the moving network..

All packets destined to MNs in a moving network are carried via the AGR. Therefore, the AGR should be placed at the optimal location considering the movement characteristics of the moving network, the location of each CN and so forth. If necessary, the AGR must be relocated. The AGR location should be chosen so that no roundabout communication paths are created between MNs to CNs as a result of network movement (factor (1)). Also, the frequency of AGR relocations should be minimized (factor (2)). If the AGR is located near the moving network, i.e., in the lower part of the core network, each communication path can be optimized and the handoff procedure can be localized (factor (3)). However, this increases the frequency of AGR relocations due to handoff. On the other hand, if the AGR is located in the higher part of the core network, the communication paths may not be optimal and the handoff procedure may

not be localized. Fortunately, AGR relocation, which is an expensive procedure, rarely occurs. As described above, there is a trade-off between factors (1)-(3). The determination of the optimal AGR location requires attention to all these factors.

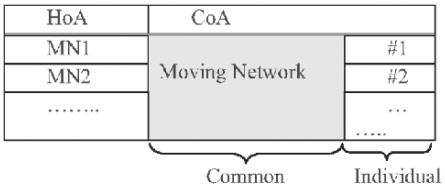


Figure 5: Hierarchical Address Management.

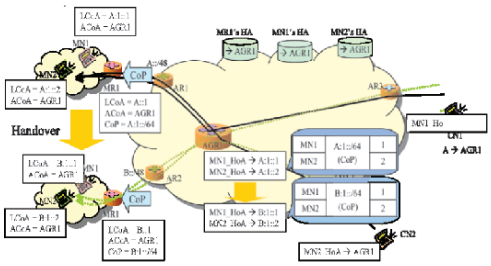


Figure 6: Care of Prefix Routing on MIP.

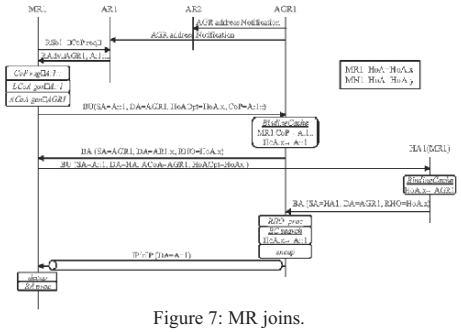


Figure 7: MR joins.

3.4 Hierarchical Address Management

Even if the techniques described in Sections 3.1 to 3.3 are used, it is still necessary to inform the AGR and the MR's HA of the updated CoAs, as all CoAs are changed when a moving network moves. The volume of handoff signals depends on the number of MNs in a moving network. Therefore, the data volume of handoff signals can become very large. To achieve seamless handoff, it is important to reduce the number of handoff signals. Hierarchical Address Management provides a solution to this problem (Fig. 5).

In Hierarchical Address Management, the CoA of each MN in a moving network is divided into the common information and the individual information. The common information indicates the location of the moving network, and this is changed when handoff occurs. On the other hand, the individual information indicates the location of each MN in a moving network, and this need not be changed even if handoff occurs. CoP, as mentioned in Section 3.1, makes this address management possible because an AR allocates an individual prefix using the same subnet mask as given to the moving network to avoid generating duplicate CoAs. The MR connected to the core network, the AGR, and the MR's HA manages the binding information using this management technique. Thus, handoff can be achieved by updating only the common information. As mentioned above, Hierarchical Address Management solves the problem by reducing handoff signal volume, not quantity.

In short, Hierarchical Address Management along with Concatenated HAs and AGR make it possible to meet requirement (3) for seamless handoff.

4 CARE-OF-PREFIX ROUTING IN A MOBILE IP NETWORK

Combining the basic techniques described in Sections 3.1 to 3.4 can yield a new routing mechanism for moving network mobility support that has the same performance as host mobility support. We call it Care-of-Prefix Routing (CoPR). Figure 6 provides an overview of CoPR. Here, the HA of each MN in the moving network holds the binding information indicating that the CoA of each MN is the AGR address. Thus, Concatenated HAs is omitted. The following details the specification of CoPR.

Figure 7 shows the sequence for connecting MR1 to AR1. AGR1 sends its address to AR1 and AR2, which are connected as the subordinate of AGR1. When MR1 connects to AR1, MR1 sends a Router Solicitation (RSol) (Johnson, 2004) containing a request for a CoP. Next, AR1 sends a Router Advertisement (RAdv) (Johnson, 2004) containing the CoP (A:1::) to MR1 and AGR1. After that, MR1 creates its on-Link CoA (LCoA) (Johnson, 2004) (A::1), sets the AGR1 address as its Alternate CoA (ACoA) (Johnson, 2004), and registers its CoP. MR1 then sends a Binding Update (BU) (Johnson, 2004) containing its LCoA and CoP to AGR1, which registers the binding information. Also, MR1 sends a BU to the HA of MR1 to register the AGR1 address as its ACoA.

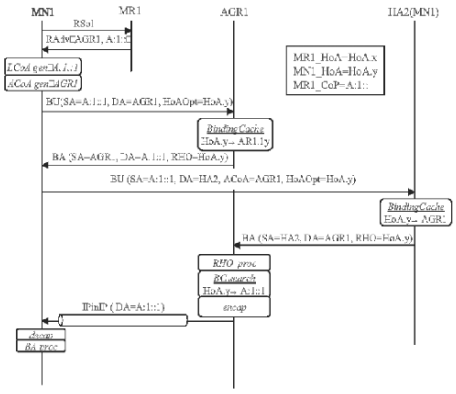


Figure 8: MN joins.

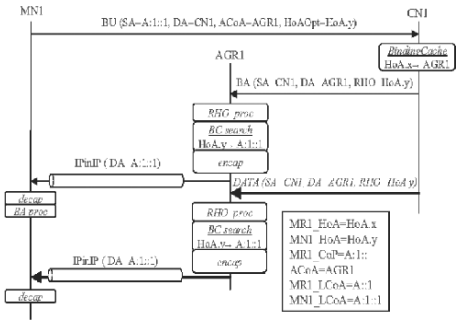


Figure 9: Optimization of the route to CN.

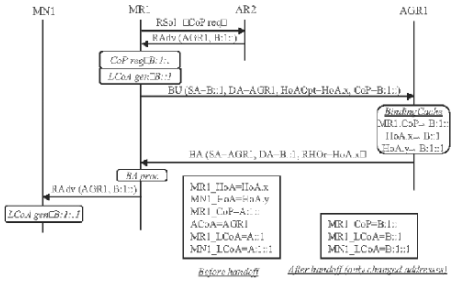
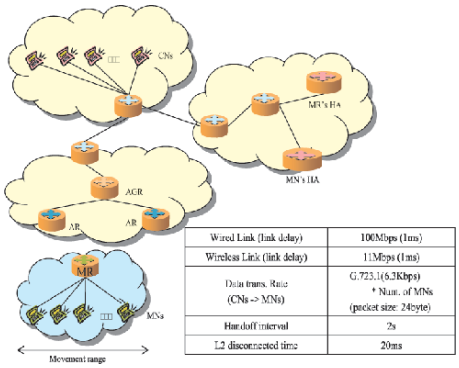


Figure 10: Handoff.

Figure 8 shows the sequence for connecting MN1 to MR1 in the case where MR1 is already connected to AR1. MN1 creates its LCoA (A:1::1) and its ACoA (AGR1 address) from the RAdv received from MR1. Next, MN1 sends a BU with its LCoA to AGR1, and a BU containing the ACoA to its HA. At that time, AGR1 caches the relation that MN-CoA1 is from the CoP of MR1.



Wired Link (link delay)	100Mbps (1ms)
Wireless Link (link delay)	11Mbps (1ms)
Data trans. Rate (CNs -> MNs)	G.723.1(6.3Kbps) * Num. of MNs (packet size: 24byte)
Handoff interval	2s
L2 disconnected time	20ms

Figure 11: Simulation Conditions.

Figure 9 shows the sequence of route optimization from CN1 to MN1. MN1 sends a BU to register the binding information that MN-CoA1 is ACoA (AGR1 address). In this situation, CN1 can send packets destined to MN1 via AGR1 using RHO. AGR1 encapsulates this packet with the LCoA (A:1::1) of MN1 after locating the LCoA (A:1::1) of MN1 in its binding cache and transmits this packet to MN1.

Figure 10 shows the sequence triggered by moving network handoff. MR1 updates its CoP and informs AGR1 of the update, after getting the new CoP (B:1::). The subnet mask of this new CoP should be the same as that of the previous CoP (A:1::). When AGR1 updates the CoP of MR1, AGR1 also updates all the LCoAs of all MNs since they are also subordinates of MR1. More precisely, only the common information is updated, since MN1 creates its new LCoA after receiving RAdv, which contains the new CoP (B:1::) sent by MR1. In CoPR, MN-CoA1 (B:1::1), which AGR1 manages, has already been updated so that it is not necessary for MN1 to send a BU to AGR1. In this way, BUs can be omitted from each MN in the moving network to the AGR. If the AGR address is changed, it is necessary to update ACoAs of MN1 and MR1.

5 PERFORMANCE EVALUATION

We have evaluated CoPR, Basic and RRH, using network simulator 2. Figure 11 shows the parameters and the topology used in the simulation.

This simulation assumed that the AGR location was optimal, as shown in Figure 11. The simulation time was 10 seconds, and the first 2 seconds were discarded to eliminate the influence of jitter. We evaluated each mechanism assuming 1, 5, 10, 100,

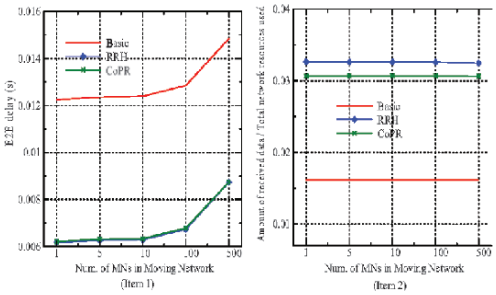


Figure 12: Comparison over Num. of MNs (Item 1).

Figure 13: Comparison over Num. of MNs (Item 2).

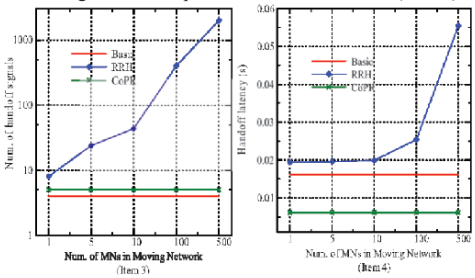


Figure 14: Comparison over Num. of MNs (Item 3).

Figure 15: Comparison over Num. of MNs (Item 4).

and 500 MNs in the moving network. The following items were evaluated:

(1) E2E delay

(2) Amount of received data/total network resources used

(3) Handoff signal overhead

(4) Handoff latency

(5) Amount of packet loss

Item 1 is the mean delay of packet transmission from a CN to an MN in the moving network, and indicates the degree of route optimization. Item 2 indicates the throughput on each hop. This should increase if the route is optimized, packet header size is minimized, and discarded packets are minimized. The inverse of this measure indicates the network resource that should be provided for given traffic. Item 3 is the number of handoff signals, i.e., RSol, RAdv, BU, and Binding Ack (BA) (Johnson, 2004), per handoff. Item 4 is the mean time from handoff initiation to completion. Item 5 is the total discarded packets caused by the handoff. Items 3 to 5 also indicate handoff performance.

Comparisons for items 1 and 2 for various numbers of MNs are shown in Figures 12 and 13.

With respect to items 1 and 2, the results of CoPR are good as shown in each figure. This is because CoPR implements both route optimization and minimization of packet header size.

With regard to item 1, CoPR is superior to Basic in terms of performance regardless of the number of MNs. The degree of superiority would increase if the HA is separated from the MR, because the packets must pass through the bidirectional tunnel from the MR to its HA. On the other hand, CoPR and RRH offer similar levels of performance since both of them optimize routing.

With regard to item 2, the ratio of CoPR performance to those of the conventional methods is almost independent of the number of MNs. The ratio is 1.90 when compared to Basic. This shows that CoPR transmits data more efficiently than Basic. This difference is due to the difference in encapsulation distance of Basic and CoPR. Basic uses a longer encapsulation distance, from the MR to its HA, whereas CoPR encapsulates only the route from the MN to the AGR. On the other hand, the performance ratio is 0.94 for RRH. The reason is as follows. In RRH, packets are transmitted with an RHO that sets two CoAs, MR-CoA and MN-CoA, from the CN to the MN. In comparison, in CoPR, the packets are transmitted with an RHO that sets one AGR address from the CN to the AGR, and also by encapsulation from the AGR to the MN. Therefore, CoPR is better and this ratio is larger if the CN is farther from the moving network than considered in this simulation environment. The reciprocal of item 2 represents the network resources needed to support the traffic of a new service. In other words, increasing item 2 makes it cheaper to put a service into operation.

Items 3 and 4 for RRH change rapidly with the number of MNs. Figures 14 and 15 show the comparisons for various numbers of MNs.

With regard to item 3, both Basic and CoPR offer low and constant values. On the other hand, in RRH, increasing the number of MNs increases the number of handoff signals. Specifically, if the number of MNs is 500, CoPR has about the same level of performance as Basic, while it requires 2,000 fewer handoff signals than RRH. The reason is that RRH demands that all MNs in the moving network send a BU to each CN and HA.

For item 4, the performance ratio of CoPR to Basic is 0.38, regardless of the number of MNs. This difference depends on the BU destination. If the HA of the MR is located farther from the moving network than considered in this simulation environment, the degree of superiority of CoPR would increase. On the other hand, the ratio of CoPR to RRH depends on the number of MNs, e.g., 0.32 with one MN, 0.11 with 500 MNs. This shows that CoPR has lower handoff latency than RRH. The superiority of CoPR over RRH is due to the fact that

the BU destination is only the AGR in CoPR, compared to all CNs and all HAs in RRH. Therefore, if the number of CNs and MNs in the moving network is increased or the distance between an MN and its HA, or between an MN and a CN is increased, the handoff latency of RRH increases dramatically. In short, CoPR is much better than RRH.

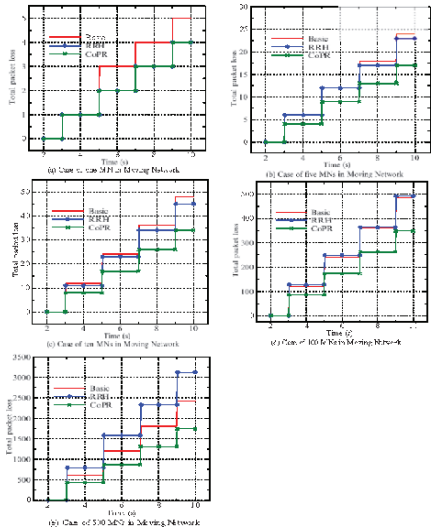


Figure 16: Comparison on each case (Item 5).

Figure 16 shows the comparisons for different numbers of MNs regarding item 5, i.e., the total packet loss. As these figures show, the amount of discarded packets on CoPR is the smallest of the three methods, regardless of the number of MNs. Additionally, the three methods have different time ranges of discarded packets. For Basic, it is from the L2 disconnect time until the binding information that the MR's HA manages is updated. For RRH, it is from the L2 disconnect time until the binding information that each CN manages is updated. For CoPR, it is from the L2 disconnect time until the binding information the AGR manages is updated. This value of RRH becomes worse than those of the other methods as the number of MNs increases. This is because the number of handoff signals increases as the number of MNs grows.

6 CONCLUSION

This paper clarified the difference between host mobility support and conventional moving network mobility support, and proposed new routing mechanisms for moving network mobility support that meet all requirements. Specifically, this paper proposed four basic techniques: Care of Prefix, which minimizes the packet header size, Concatenated HAs and Hierarchical Address Management, which reduce the number and volume of handoff signals, Aggregate Router, which aggregates and localizes handoff signals, and CoPR, which is a mechanism for applying these basic techniques to MIP.

We verified the effectiveness of our proposed mechanisms using network simulator 2. Quantitative analyses showed that CoPR is the best in terms of five measures: E2E delay, amount of effective received data / total used network resources, amount of handoff signals, handoff latency, and amount of discarded packets. As mentioned above, CoPR is superior to the conventional solutions proposed in NEMO. We will construct an experimental system and verify the feasibility of the proposal mechanisms.

ACKNOWLEDGMENTS

The authors would like to thank Mr. Yoshizawa and Mr. Fukazawa of NTT COMWARE Corporation for their useful advice on the simulation.

REFERENCES

IETF, Available: http://www.ietf.org/

Yumiba, H., Imai, K. & Yabusaki, M. 2001. *IP-Based IMT Network Platform*, In IEEE Personal Communication Magazine, pp. 18-23

Johnson, D., Perkins, C. & Arkko, J. 2004. *Mobility Support in IPv6*, RFC3775 *NEMO*, Available: http://www.mobilenetworks.org/nemo/*The Network Simulator*, Available: http://www.isi.edu/nsnam/ns/

Devarapalli, V., Wakikawa, R., Petrescu, A. & Thubert, P. 2004. *Network Mobility (NEMO) Basic Support Protocol*, Internet Draft: draft-ietf-nemo-basic-support-03.txt

Thubert, P. & Molteni, M. 2004. *IPv6 Reverse Routing Header and its application to Mobile Networks*, Internet Draft: draft-thubert-nemo-reverse-routing-header-05.txt

Deering, S. & Hinden, R. 1998. *Internet Protocol, Version 6 (IPv6) specification*, RFC2460

Suzuki, T., Hiyama, S., Igarashi, K., Kawakami, S. & Hirata, S. 2003. *Routing Management for Moving Network*, In IEICE General Conference.

SERVICE INTEGRATION BETWEEN WIRELESS SYSTEMS
A core-level approach to internetworking

Paulo Pinto and Luis Bernardo
Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, P-2829-516 Caparica, Portugal
Email: pfp@uninova.pt, lflb@uninova.pt

Pedro Sobral
Faculdade de Ciência e Tecnologia, Universidade Fernando Pessoa ,Porto, Portugal
Email:pmsobral@ufp.pt

Keywords: Interworking of wireless systems, wireless computing, handovers, mobility management.

Abstract: The greater bandwidth provided by wireless LANs can be a precious asset to the wireless ubiquitous computing if the integration with 3GPP systems is done at a certain level. This paper presents a proposal to integrate wireless systems at core network level. Service integration becomes very powerful and easy. The system is not so dependent on the critical latency of vertical handovers and the users feel a unique system providing services. Little changes are required to the current 3GPP core network. Our architecture uses the GPRS as the primary network and integrates WLANs as secondary networks, used on an availability basis. Sessions on secondary networks survive disconnection periods contributing to a seamless service provision to the user. The paper describes the overall architecture, the changes that are needed at the current 3GPP core, and the operation of the secondary networks on the aspects of data routing and security associations. Highlights about the application model are presented at the end.

1 INTRODUCTION

Cellular systems like Global System for Mobile Communication/General Packet Radio Service (GSM/GPRS), and its successor UMTS (Universal Mobile Telecommunication System) already provide IP-services in a ubiquitous mode. However, there are obvious limitations on bandwidth due to their coverage requirements. Wireless LANs (WLANs) have been seen as a useful add-on to provide islands of greater resources. Ways to integrate these systems are being developed and a challenge discussed in this paper is how the integration can be done to allow new services to appear (e.g. exploring mobility) and still support the existing ones.

Current proposals either envisage a complete integration of WLANs in the cellular system's architecture (*tightly coupled*) or provide integration in such a way that the systems interact poorly (*loosely coupled*). The former solution does not enlarge the type of services that can be designed from the current cellular architecture framework and the latter makes it difficult to provide a real sense of service integration amongst the various access networks.

This paper presents a solution based on an integration performed at core network level amongst the major components (such as the SGSN – Serving GPRS Supporting Node). The system can be used in different business scenarios and not only in a 3GPP operator owned WLAN infrastructure. The paper starts with a scenario of a service to justify why a new approach to integration is viable. Our approach maintains a user session regardless of the precise Radio Access Network (RAN) used, claims that vertical handovers (between RANs) are not needed, and RAN switch can better be performed at core network level (in the UMTS sense).

Each UE (user equipment) can maintain at least one IP session over a specific RAN and can always use the session, even when it is outside the coverage area of that RAN (by using other active RANs). Services can access core-level information to: (a) improve the way they use the communication links to the UE; and (b) handle and adapt to UE mobility and connected periods (when the UE is inside the coverage of a WLAN). In order to implement our system only minor (software) modifications to the current 3G core network need to be done.

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 193–200.

2 ENVISAGED SCENARIO

Initial assumptions – First, the reality today is that the cellular network is ubiquitous, covering 100% of the populated areas. It is very unlikely that any other radio network system will have such coverage. The consequence is that any other network will have dark areas, and supporting users in these networks alone is not feasible. Second, our UEs are equipped with two (or more) wireless interfaces working simultaneously. Third, WLANs can be owned by private organizations with agreements to the 3GPP system operators or owned by the operators themselves. Fourth, the security control provided by the USIM smart cards and global roaming agreements between 3GPP system operators form the largest operational security system in the world to date. AAA (Authentication, Authorization and Accounting) procedures between 3GPP systems and WLANs are on the verge of being approved (3GPP, 2003) and we assume them in our system.

Down load service– Our service example extends the infostation model presented in (Frenkiel, 2000) with cellular network integration (the real subject of the example – download of data – could be part of a more sophisticated application).

A user is at home and uses the GPRS interface to start a service to download some bulk data. In his way to work, the system will try to use the WLAN RAN (near semaphores, etc.) to deliver the data. Eventually, all the data will be transferred.

In the rest of the paper, we consider GPRS as a packet service in both 2.5G and 3G systems standardized by 3GPP.

3 SYSTEM OVERVIEW

The capacity of radio cells will increase in the future. Cells will be smaller than the current ones. As stated in (Frodigh, 2001) we also agree that extremely high rates will not be necessary everywhere, but just in small hotspots. The question is how to integrate these hotspots?

3.1 Hotspot Integration

One possibility is that they are part of the cellular network as an ordinary cell. The network would predict the user movement (using the cell information) and could schedule the sending of large bulks of data when a hotspot becomes available. However, implementing such a facility at network level can be rather complex (as there is not enough relevant information). Moreover, unless applications have knowledge of the differences in cells and adjust to specific cell data rate conditions a user might experience lack of bandwidth just because he stepped out of the coverage.

Another possibility is that high bandwidth cells are seen as special cells, not integrated in the cellular system and having a special (direct) connection to a packet data network. The user knows he is using a different interface and stepping out of coverage is easy to detect.

There are proposals for WLAN integration covering both possibilities. The *tightly coupling* option (Salkintzis, 2002) state that cells should be integrated at a low level offering an interface compatible with the 3GPP protocols. Besides the drawbacks listed above there are still the following disadvantages: (a) the WLAN must be owned by the 3GPP operator (to avoid strong exposure of core network interfaces); (b) cell displacement and configuration demands carefully engineered network planning tools and WLAN integration becomes difficult. Moreover, a great deal of control procedures are based on configuration parameters (CellID, UTRAN Registration Area (URA), Routing Area (RA), etc.) and WLAN cells have to comply with them; and (c) paging procedures and handovers (including vertical handovers) have to be defined and some technologies (e.g. IEEE 802.11) are not so optimized to make them fast enough.

The loosely coupled option (Salkintzis, 2002 and Buddhikot, 2003) assumes there is a WLAN gateway on the WLAN network (with functionalities of Foreign Agent, firewall, AAA relaying, and billing) and the connection to the 3GPP core is via GGSN (Gateway GPRS Support Node) (with a Home Agent functionality). It only makes sense to use this option with dual-mode UEs because a vertical handover to WLANs would disconnect the UE from all the functionality of the cellular networks (paging, etc.). One advantage is that high-speed traffic is never injected into the 3GPP core network. A major disadvantage is the degree of integration. WLAN networks are handled independently and will be used on an availability basis by the users, whom have to stay within the same coverage. Any service provided by the 3GPP (SMS (Short message Service), etc.) has to consider the cellular system's internet interface. Any exploitation of the UE's mobility (both in the cellular system and inside the WLAN island) is hidden by the mechanism of Mobile IP, for instance. From the applications point of view, the UE is stationary placed inside a big cloud called GPRS (or WLAN). I.e. it has a stable IP address and any mobility inside the 3GPP network is not seen from the exterior.

3GPP (3GPP, 2002) defined six scenarios of increasing levels of integration between 3GPP

systems and WLANs. Scenario 3 addresses access to 3GPP PS services and includes access control and charging. (3GPP, 2003) specifies how it should be done. A loosely coupled approach was adopted but the data routing aspects are still not fully agreed (the specification covers mostly the access control and charging).

Our proposal for hotspot integration is somewhere in between the tightly and loosely options – it is at core network level. It allows the use of WLANs as a complement to the GPRS network. It is not fully incompatible with the 3GPP effort, as it will be described below.

3.2 Primary and Secondary Networks

In our system the GPRS network is the glue for all the other RANs. It is the primary network having all the control services (paging, etc.). All secondary networks become simpler and can have control services of their own not seen at core level (i.e. they are simply internal optimizations). Almost all of the works in internetworking assume that all these features (including paging) exist in all networks and are seen at core level. IDMP (Misra, 2001) is one of the exceptions stating that they should be customized. The most similar approach to ours was taken by MIRAI (Wu, 2002). Their primary network is a collection of BANs (Basic Access Network). Each BAN contains the usual control services, and is controlled by a CCN (Common Core Network) manager. A user selects a RAN based on a list provided by the BAN considering user location and preferences. Although, the authors consider a long list of issues to help the UE choose the RAN, some too low level or "external" reasons (e.g. battery life) can lead to unexpected choices from the applications' point of view. CCN handles micro-mobility (possibly inter-RAN) and participates in macro-mobility. The control features of the BAN are very similar to the ones in UMTS. It could have been implemented by the 3G system (as also stated in (Wu, 2002)) but MIRAI authors decided to implement a new radio interface.

In our system a WLAN RAN is a set of islands. Each island is formed by a set of cells and is controlled by an Island Manager (IM). Islands do not cover the entire space (i.e. there will be dark areas). All islands of a certain technology are seen by the primary network as a ***Hotspot Network*** (HN) – a secondary network.

A user is connected to the GPRS network and can have other sessions simultaneously. Each HN supports the notion of a session (i.e. IEEE 802.11 has one, HiperLan has another, etc.). Differently from the other proposals is the fact that a session survives the disconnection periods when the user is moving in a dark area of a certain WLAN. For instance, the user began a 802.11 session at the airport, took a taxi to a hotel, and when he is in the hotel, the same session is still on using the WLAN infrastructure of the hotel (it is assumed that both have agreements with 3GPP operators). On the way from the airport to the hotel, if the user needs to be contacted in the context of that session the primary network is used.

Other works consider all RANs at the same level. (Tönjes, 2002) defines a flow router at the core that uses all RANs. This will lead to the existence of control functions in all of them. If only one is chosen to have these features the system will fall back to ours. Moreover, with a monolithic core it would be more difficult to add a new RAN.

4 ARCHITECTURE

Figure 1 shows the architecture for the data traffic (no access control, billing, etc.). The new components are

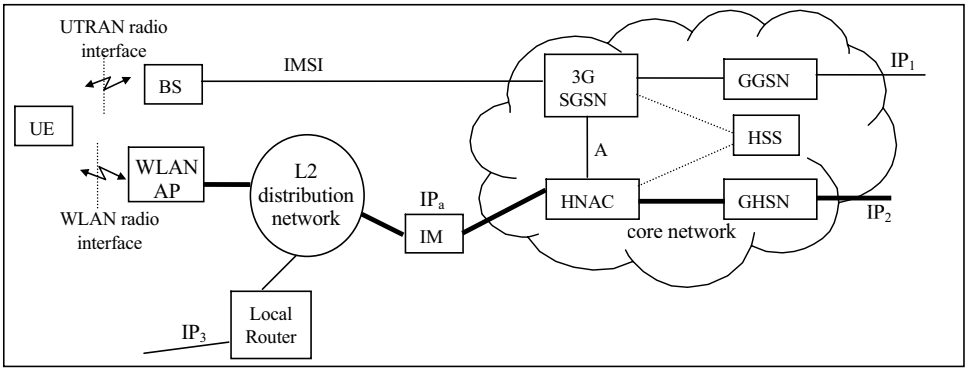


Figure 1: Data traffic in the hybrid network.

the HNAC (Hotspot Network Area Controller) which controls one (or more) island, and the GHSN (Gateway Hotspot network Support Node) which is responsible for context management and Internet access. The thicker lines belong to the core but they are not present in the current 3G core. All the high speed traffic goes through them not overloading the current 3G infrastructure.

The 3GPP specification for scenario 3 (3GPP, 2003) has a component that merges the HNAC and the GHSN, called PDG (Packet Data Gateway). The PDG is not connected to the SGSN (line A) as we propose and all data integration between the systems is done at IP level.

An UE has its identification at core level in the form of an IMSI (International Mobile Subscriber Identity) and the attachment procedure for GPRS is the standard one (with temporary identifiers). In the GPRS world an IP session can be established via the GGSN (PDP context), having a routable address, called IP_1.

If the UE senses a WLAN to which it can perform a connection establishment, it does so. From that time on it can use the 'local router' to access the Internet directly. An IP_3 address is used for that path (whether this address is a care-of-address and whether the local routing is performed at level 3 or level 2 is irrelevant to this paper). If the WLAN has roaming agreements with a 3GPP operator the UE can perform an attachment procedure with the 3GPP operator (3GPP, 2003). The attachment defines a *local identifier* at core level for the UE in that WLAN (possibly with temporary addresses, too). In figure 1 an IP address was used as an example for the local identifier (IPa). From this time on an UE identified by the IMSI can be contacted via UTRAN using the IMSI, or via WLAN using the IPa. It is important to note that no assumption is made about an *all-IP* technology in our system. It is sufficient that it is *IP-enabled*. I.e. the UE communicates at IP level with the core, but the core forwards packets to the IM to be delivered to an UE with a specific local identifier. The core does not assume any delivery protocol in the island. The IM can use a layer 2 routing if it suits better. The important thing is that IPa is stable. If the UE wants to use the Internet via the WLAN it creates a PDP context (in similar modes as to the GPRS case) and a routable address IP_2 is defined at GHSN. Every time there is an attach update (in a different WLAN, for instance) a new IPa is chosen but IP_2 remains the same.

IP_1 is the main, fixed, UE address. IP_2 and IP_3 should be used on a temporary basis (e.g. client applications). Therefore, reuse of addresses can be made making the system scalable.

4.1 Overview of the Interactions

The HSS (Home Subscriber Server) has the operational information about the UEs. Besides the GPRS-related parameters that the HSS already has, there is the information if an UE is HN attached, has a session established and if it is currently inside a WLAN coverage area (and the identification of the HNAC responsible for it). SGSN and HNAC will go to HSS to get updated information. The HSS also provides authentication vectors, subscriber profiles, and charging information.

The communication between the core (SGSN and HNAC) and the UE can use any RAN. We will describe two approaches: the first one, the smooth transition, consists in keeping the GPRS almost as it is with little add-ons. Any PS traffic will use UTRAN but the HNAC can communicate directly with the UE via WLAN, or can relay the traffic through the SGSN to be delivered to the UE via UTRAN (using the interface link A in figure 1. A more concrete description is given below); the second approach is more abrupt – both SGSN and HNAC can convey their traffic through the other component if they see some advantage. Currently, a tunnel called GTP-U (GPRS Tunneling Protocol – for User Plan) is established between the GGSN and the serving RNC (Radio Network Controller). In our second approach the tunnel goes as far as the SGSN and a new tunnel is formed from there on. It is a return to the original GPRS specification. Figure 2 shows the protocol stack at an UE. There is a Connection Manager (CM) that manages the status of both connections and offers a unique interface to both RANs. The Delivery Service is a confirmed service and switches to the UTRAN if it senses a failure in the WLAN. If more

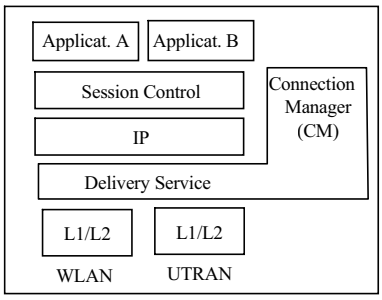


Figure 2: Protocol stack at a UE.

than one RAN is active the default one for each message is used. The CM can signal the applications (or be queried by them) about the current status of a

specific connection. With this information, applications can avoid using the link if the proper interface is not active (transferring only urgent information, for instance). The CM is able to contact each of the core network components (SGSN or HNAC) either directly or via the other RAN (for link maintenance messages, etc.). The Session Control is responsible for session survival when the UE is in dark areas.

For the smooth approach the following interactions are needed: (a) permission by the SGSN to create a session between the UE and the HNAC (using a PDP context just as the GGSN uses them, with the Session Management Protocol) (Kaaranen, 2001). A GTP-U tunnel is created between the HNAC and the UE (more precisely with the serving RNC); (b) an event service from the SGSN to notify relevant events – "UE availability", "cell update", "routing area update", "positive cell identification" and "undefined cell identification". It is important for session management by the HNAC; and (c) mobility management information by the SGSN (cell identification if in GPRS state *ready*, or routing area identification, otherwise) - it can be useful for the HNAC. Suppose HNAC has a relation between cells and WLAN islands topologies. It can force the WLAN interface to switch off if no islands are known in a certain routing area, for instance. It is also important because HNAC change of responsibility can happen when the UE performs a routing area update.

For the abrupt approach the current GTP-U tunnels have to be divided in two parts: one from the GGSN to the SGSN; and another from the SGSN to the RNC. The same will happen from GHSN to HNAC, and from HNAC to IM. This separation allows the second tunnels to be established either via the default RAN, or via the other RAN.

4.2 Scope of Integration

In our system there is no need for vertical handovers because the GPRS session is always on and the other RANs are used as a complement. Communication to the UE can use indistinguishably any available RAN. A total switch of the communication from one RAN to another is performed by the core components, so no information is ever lost. In systems with traditional vertical handoff, the dominant factor is the time the UE takes to discover that it has moved in/out of coverage (i.e. the cell has to become active or inactive) (Steem, 1998). Using RANs in a complementary form as we do, this time is not so critical, and the GPRS can provide a minimal bandwidth.

As the integration is performed at core level current services can work with secondary RANs in a very easy way. Figure 3 (taken from (3GPP, 2003)) shows how 3GPP plans to support SMS over WLAN. A service specific gateway, IP-SM-GW, must exist and offer an interface similar to an MSC or an SGSN (interfaces E or Gd) to the GMSC/SMS-IWMSC. The address of this gateway is returned by the HSS in the "send routing information for short message". This gateway has a private database to associate MSISDN to IP addresses. UEs in WLAN have to specifically register and specifically authenticate for SMS services and have secure associations to the gateway. The gateway communicates with the UE via Internet.

In our system (abrupt approach), the SMS service could be provided without any modification. The SGSN just gets the message and can use the HNAC to convey the message to the UE, using the secure associations that are already in place.

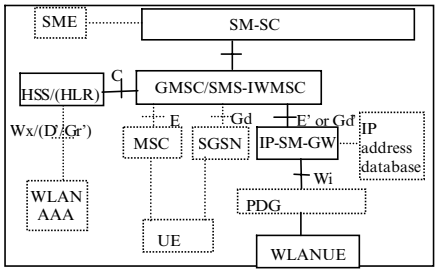


Figure 3: Support of SMS over WLAN.

4.3 Application Support at Core Level

HNACs have already the task of maintaining sessions between appearances of the UE in WLAN islands. A step further is their ability to work with the applications in order to take advantage of the mobility (and connect times) of the UEs to perform the application task in a specific manner. This is not the traditional approach in the Telecom world and resembles more the activity of a middleware level managing mobility.

In the Telecom world networks are seen as closed systems that offer services. Services are carefully specified procedures that use lower level procedures called *bearer services*. The control procedures of the network are seldom accessible from the exterior and well protected from external components. It is interesting to see that the same approach is being planned for the introduction of Voice over IP (VoIP) services (Lin, 2002) on top of GPRS. The RAN and the GPRS network together are called the *bearer network*. Through the *Gm interface* (which includes

radio, Iu, Gn, and Gi), the *bearer network* provides bearers for signaling (control plane) and data (user plane) between the UE and the IP Multimedia Subsystem (IMS) (placed outside the core network). The *bearer network* nodes (RAN, SGSN, and GGSN) are not aware of the multimedia signaling between the UE and the IMS.

The typical way to add new functionalities and behaviors to networks is using frameworks such as Intelligent Networks, or in case of UMTS, the CAMEL (Customized Applications for Mobile network Enhanced Logic), (Kaaranen, 2001). However, the extensions are traditionally related with the basic services and with inter mobile-network interactions. For instance, personalization of services, different control over switched circuits, virtual home environments when roaming, etc.

In our system, HNACs can have a standard (and protected) programming interface to be used by third-party organizations to build services and applications that take advantage of the information gathered at the core (and not accessible today). The end of this paper has an example of one.

5 HN OPERATION

The interaction between an island and its controlling HNAC is performed by the *Island Manager* (IM). The IM provides a stable identifier for the UE and forwards packets between the UE and the HNAC. In terms of security the 3GPP network can offer an authentication service to the WLAN owner (WLAN connection establishment is not covered by the 3GPP

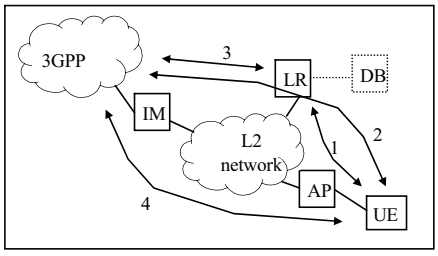


Figure 4: Security flows for UEs.

specification, obviously). It is important that both networks rely very little on each other (not disclosing authentication vectors, for instance). Figure 4 shows the proposed setup. It consists of two, almost similar, phases. The first provides WLAN authentication assisted by the 3GPP network, and the second provides 3GPP authentication via WLAN. The UE senses a WLAN and creates a provisional secure association with the local router (LR) (1) (it is assumed that the AAA functionality is inside the LR). Using this association it sends a message to the LR to state its willingness to authenticate. The LR triggers an authentication process within the HNAC. The HNAC gets authorization vectors from the HSS and issues a challenge. The local router relays the challenge and the corresponding response between the HNAC and the UE (using, for instance EAP Response/Identity) (2). The result of the authentication is given to the LR by the HNAC (EAP-Success/EAP-Failure) (3). At this moment the LR knows the UE has the identity it claims it has. The LR checks if the UE can use the WLAN, by consulting a local database of users. If so, it creates a definitive secure association (in the scope of WLAN), provides the keying material to the UE for local WLAN use, and informs the address of the IM. The 3GPP could also approve a user not belonging to the local WLAN community, in which case the LR will tell the user that a local session cannot be established but the IM address is given for a WLAN-3GPP session.

If the UE has passed the first phase, it can now start an authentication process with the 3GPP to create a context there (4). The secure association is created with the 3GPP without intervention, or knowledge, of the WLAN. The attachment to the 3GPP network is covered in (3GPP, 2003) and uses the EAP authentication procedure, providing enough keying material for a secure tunnel to the PDG (or HNAC in our case) through the IM. Once attached, the HSS has the indication that the UE exists and a session can be established (both by the UE and by the 3GPP). A WLAN session can also be established using the UTRAN interface (particularly useful in dark areas). Each time a new island is entered a fast update procedure must be done.

The concrete mobility management protocol used inside the island and any mechanism to save power or bandwidth are irrelevant and should not be seen from the 3GPP. I.e., a micro-mobility move must not change the local (IPa) address to maintain the secure associations and the information in the core. Any possible paging mechanism prior to the delivery of a message is also hidden from the 3GPP. From the core level point of view a packet is simply delivered.

Figure 5 shows the state diagram of the interaction UE-3GPP. It is assumed that the UE is always attached to the GPRS. Its idle state is the *Disconnected* state – there is no operational WLAN information about the UE in the core.

When the UE attaches and creates a PDP context, the address IP_2 is defined, and the HSS has information about its existence. An HNAC will be responsible for it and the state changes to *Registered*. In this state it is assumed that the UE is always reachable via IP_2.

Any communication to/from the UE is done in the context of a *session*. A session represents extra attention by the HNAC to the position of the UE (using also the mechanisms provided by the SGSN as stated above). The *Session* state is entered when either the UE or the HNAC issue a *Start of Session* message.

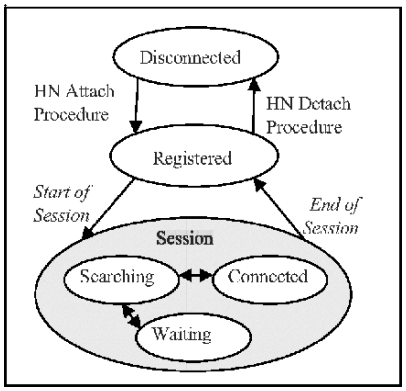


Figure 5: UE-3GPP State Diagram
UE-3GPP State Diagram.

The sub-state *Connected* means that the UE is inside an island. The sub-state *Searching* is entered when the UE is out of coverage. In the *Searching* sub-state, the HNAC forces the UE to be in active GPRS state to know its cell ID. It can happen that the UE is in a cell that has no islands nearby. In this case the HNAC can order the UE (via UTRAN) to go to sub-state *Waiting* to save battery power. When the UE moves to a cell where an island exist it is told to change again to *Searching* (note that these sub-states are simply optimizations and can exist, or not).

Depending on the application it can be easy to know when a session finishes, or not. If it is, an *End of Session* message is sent. If it is not, a watchdog mechanism based on inactivity triggers the sending of the message, for instance.

6 APPLICATION MODEL

Applications may interact with external networks using one of the three connections: UTRAN (IP_1), WLAN direct (IP_3) or WLAN-HN (IP_2). The application models for the first two follow the traditional Internet models – correspondent nodes communicate with fixed remote IP addresses and send data as soon as it is available. For the WLAN-HN we can have a different approach that optimizes the use of scattered hotspots over a ubiquitous 3GPP network. With proper support at the core network, these applications will be able to maintain sessions independently of the hotspot availability, and communicate the bulk part of the data only when the UE is in the *Connected* sub-state.

Application functionality is divided between the UE and the serving HNAC (figure 6). In the UE we have a *front-end component* implementing the user interface and interacting with the session control entity and the associated lower services (which behave as a middleware platform for the applications). In the serving HNAC we have the *back-end component* that cooperates with its peer on the UE and maintains a stable interface with external entities. By stable it is meant that any optimization use of the air interface is hidden from the external applications. These components work in the context of the IP_2 session (Fig. 1), but can communicate with each other using the WLAN RAN or the UTRAN. They can work in an "*Always Connected*" mode, using the SGSN each time the UE is in an HN dark area, or, more interestingly, in "*Hotspot Connected*" mode, communicating only urgent information via SGSN while waiting for the UE to become *Connected* again.

The middleware performs session and mobility management providing applications with context information. The middleware layer gathers UE mobility information from the SGSN and network status from probing its network interfaces. These events are used by the middleware services to update the execution environment parameters. Using this information, applications are able to adapt their behavior to different network conditions and mobility scenarios.

For instance, if a user wants to download a news summary stored in a web site, both components start a new session, and the *back-end component* will start to fetch the videos. If the UE gets out of coverage the *back-end component* can store a portion of the data waiting for the UE to pop up. Later, when the UE enters into a hotspot, the information will be forwarded to the *front-end component*. In the meantime both modules can exchange control information via the SGSN/UTRAN. This pre-fetching feature optimizes UE connection time with HN, avoiding fetching delays from exterior networks. The back-end context has to be highly mobile because it might have to change to another HNAC pursuing the UE (dashed arrow in the figure). Information can be stored either in the HNAC or in a server close to the core network with a guaranteed delay for access (avoiding copying when the serving HNAC changes).

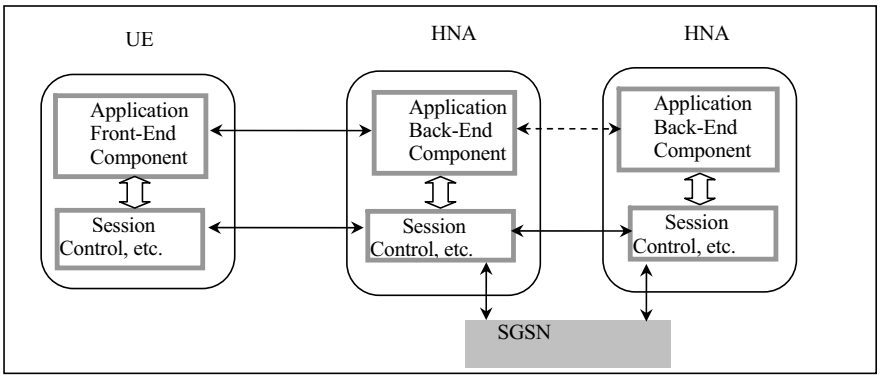


Figure 6: Functional Blocks for HN session applications.

7 CONCLUSIONS

The internetworking of wireless infrastructures performed at core level with a pivot network seems a simple and executable model. First, as most of the control features already exist in the PLMN, they can be absent in other networks. Second, because certain details on secondary networks (such as micro-mobility) are not managed at core level. Third, because it defines an environment where new features and services can be added to the core.

The addition of new modules at core level with standard (and protected) programming interfaces can open up new possibilities to explore terminal mobility (a topic that is absent today).

Our solution does not impose relevant requirements to the overall system: the architecture does not need to be all-IP; there is no critical dependence on vertical handovers; and, does not create extra load to the current 3GPP core network.

Topics that are relevant for further work include the algorithms to be used on top of the HNACs to explore the mobility of UEs and their connection periods, and the viability of service continuity using this type of handovers.

REFERENCES

3GPP, 2003. Group Services and System Aspects; 3GPP system to WLAN internetworking; System Description (Release 6), TS 23.234.v.2.3.0, Nov.

Frenkiel, R., et al., 2000. The Infostations Challenge: Balancing Cost and Ubiquity in Delivering Wireless Data, *IEEE Personal Communications*, v. 7, pp. 66-71, n.2, April.

Frodigh, M., et al., 2001. Future-Generation Wireless Networks, *IEEE Personal Communications*, v 8, pp. 10-17, October.

Salkintzis, A., et al., 2002. WLAN-GPRS Integration for Next-Generation Mobile Data Networks, IEEE Wireless Communication, v.9, pp. 112-124, October.

Buddhikot, M., et al., 2003. Design and Implementation of a WLAN/CDMA2000 Interworking Architecture, IEEE Comm. Mag., pp. 90-100, November.

3GPP, 2002. Group Services and System Aspects Feasibility study on 3GPP system to WLAN internetworking; (Release 6), TS 22.934.v.6.1.0, Dec.

Wu, G., 2002. MIRAI Architecture for Heterogeneou Network, IEEE Comm. Mag., pp. 126-134, February. s

Tönjes, R., 2002. Flow-control for multi-access systems, PIMRC'02, pp. 535-539, September.

Steem, M., Katz, R., 1998. Vertical Handoffs in wireless overlay networks, Mobile Networks and Applications, v.3, pp. 335-350.

Lin, Y., et al., 2002. An All-IP Approach for UMTS Third-Generation Mobile Networks, IEEE Network, v.16, n.5, pp. 8-19 September/October

SPUR: A SECURED PROTOCOL FOR UMTS REGISTRATION

Manel Abdelkader and Noureddine Boudriga
National Digital Certification Agency
3 bis rue d'Angleterre, Tunis RP 1000,Tunisia
Email: {maa,nab}@certification.tn

Keywords: UMTS Release5, Registration, Authentication, IMS Security , SIP Security, Security Associations

Abstract: This paper presents a new scheme for mobile identification and registration in UMTS networks. Our approach attempts to alleviate different limitations observed with the current solutions (such as the 3GPP). It guarantees the protection of the data transmitted on the SIP messages during the registration procedure. Our method provides the authentication of the main entities involved in the registration procedure. It develops a mechanism for the management of relating security associations.

1 INTRODUCTION

Recently, the development of the Universal Mobile Telecommunications System (UMTS) architecture has known great evolutions as it can be noticed with the 3GPP specifications (Kaaranen et al., 2001). Since its release 5, the UMTS network has emerged to an all IP network leading to the introduction of new protocols and procedures (TS 23.228, 2003; TS 22.228, 2002). Among the most important subjects that have been discussed for the all IP network, one can find the problem of how to overcome the different threats applicable to the UMTS networks (TS 33.900, 2000; TS 33.120, 2000; TS 21.133, 2001).

The registration procedure of a mobile to a service provided by a UMTS network represents one of the critical phases that should be protected. During this phase, there is no fixed definition of the mechanism that allows to protect the integrity, confidentiality and authentication of the Signaling Initiation Protocol (SIP) messages involved with the Internet Multimedia Subsystem Authentication and Key Agreement (IMS AKA) process (TS 33.203, 2002; TS 24.29, 2002; Rosenberg et al., 2003).

Different proposals have been presented to provide registration ((S3-000689, 2000) and (S3z000010, 2000)). Authors of (S3-000689, 2000) have proposed that the Proxy Call Session Control Function (PCSCF) performs the IMS AKA with the Mobile Station (MS) and terminates integrity and confidentiality protection of the SIP messages transmitted by MS. However, the protection of the remaining segments of the communication toward the Serving CSCF (SCSCF) is based on the network domain features using Internet Protocol Security (IPsec). Therefore, the SCSCF may not be able to authenticate users at the service level. Authors of (S3z000010, 2000), on the other hand, have proposed that authentication and re-authentication procedures should be made by the Home Subscriber Server (HSS) using AKA process. The integrity and the confidentiality keys are then transmitted to the SCSCF and the PCSCF to insure the protection of the SIP messages. The main drawback of this approach is the important load added to the HSS. The 3GPP scheme allows to overcome some drawbacks of the previous two proposals by moving the authentication process to the SCSCF.

The IMS AKA (TS 33.203, 2002) presents itself other lacks of security, which include for example the following facts: (1) it transmits (in clear) the mobile private data; (2) it does not provide the authentication of the serving network to the user; and (3) it allows the SCSCF to attribute the user private keys to the PCSCF, which reduces the user's level of security. Limitations can be at the origin of different attacks such as masquerading and man in the middle.

In this paper, we propose a secured registration procedure, called *SPUR scheme*, in all IP UMTS networks that overcomes the previous mentioned limitations. We will mainly focus on the registration of a mobile to a service and provide protection schemes of the data transmitted on the SIP messages. Our method

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 201–209.

is independent from the security mechanisms adopted in the lower layers. We also develop in this paper a proposition for a secured management of the security associations that we define for need of protecting the communication between the mobile and the IMS.

The remaining part of this paper is organized as follows: Section 2 develops the SPUR scheme and describes all its steps. It also presents a procedure for re-registration. Section 3 adapts the concept of security association to protect the security elements needed for the execution of SPUR. A secured model for security associations management is also defined. Section 4 analyzes SPUR's features and compares it to 3GPP. Section 5 develops a SPUR simulation, where the effects of message size on the error probability and the additional flow between nodes are estimated. Section 6 gives the conclusion of this paper.

2 THE SECURED PROTOCOL FOR UMTS REGISTRATION

The secured protocol for UMTS registration (SPUR) is designed to increase the security level of the registration process in the UMTS networks. It adds different security measures to the registration protocol as adopted by the 3GPP. It includes two procedures: the initial registration and the re-registration procedures. The following subsections develop these procedures.

2.1 Terms and Notations

Nodes of the IMS subsystem contribute to the accomplishment of SPUR. However, for sake of simplification, the most important entities involved with SPUR are the following:

• The PCSCF: The Proxy Call Session Control Function behaves like a proxy. It accepts the MS requests, serves them internally or transfers them. In the case of registration, the PCSCF transfers the SIP REGISTER request of a user to an I-CSCF according to the home network domain name of the MS(TS 24.229, 2002).

• The ICSCF: The Interrogating Call Session Control Function is the contact point within an operator's network for all connections related to subscribers of this network. In the case of registration, upon the receipt of SIP REGISTER request, the ICSCF gets the address of the SCSCF from the HSS(TS 24.229, 2002).

• The SCSCF: The Serving Call Session Control Function acts as a SIP registrar. It provides services to the MS and controls the sessions of the users(TS 24.229, 2002).

• The HSS: The Home Subscriber Server is the master database for users containing their subscription related information(TS 23.228, 2003).

The terms used in the sequel by SPUR scheme are the following:

• IMPI, IMPU: the private and public identity of a user.

• K_{PX}, k_{pX}: the public key and the private key of X.

• $Cert_X$, $Cert_{HP}$, $Cert_{HS}$: the relative identity Certificate of X and of the PCSCF delivered by the HSS, and the attribute Certificate of the SCSCF delivered by the HSS.

• ID_X: the identifier of X(could be an IPv6 address).

• AK: authentication key shared between MS and HSS(TS 33.102, 2000).

• K_{si}: the session key established between the PCSCF and the MS.

• req_i, res_i: challenge and response used between the MS and the PCSCF during the establishment of K_{si}.

• AV_i: the authentication vector number i of a user as defined in(TS 33.102, 2000).

2.2 The Secured Registration Protocol

The registration procedure is initiated by a mobile when it wants to access a service, for the first time. The deployment of SPUR scheme supposes the satisfaction of the following assumptions:

• Every mobile MS possesses an identity certificate delivered by its home network.

• Every node of the Internet multimedia subsystem has an identity certificate. This includes the PCSCF, ICSCF, HSS, and SCSCF.

• The different Certification Authorities (CA) serving the function of publishing certificates of the above mentioned entities are linked to a Bridge Certification Authority (BCA)(Hastings and Polk, 2000) to ensure cross-certification.

• The signaling protocol used between the nodes of the IMS is assumed to be the SIP-EAP-TLS.

The registration protocol (as depicted by Figure 1) is a 15-step procedure defined as follows:

Step1. Mobile MS signs its private identity with its private key (k_{pMS}) and encrypts it with the public key of its HSS (K_{PHSS}). After that, the MS sends message M_1 to the related proxy PCSCF

$$M_1 = \{E = [(IMPI)\_k_{pMS}]\_K_{PHSS}, Cert_{MS},$$

$$Cert_{HSS}, IMPU\}$$

(where (-)_k stems for the encryption function using key k). MS can obtain the address of the PCSCF from the Gateway GPRS Support Node after the success of a PDP ATTACH process (TS 23.060, 2002).

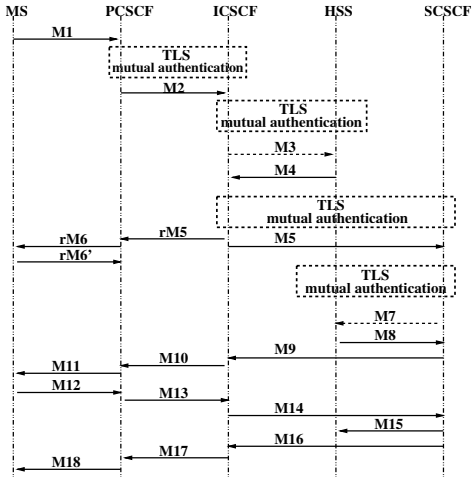


Figure 1: SPUR Architecture.

Step 2. Upon receipt of M_1, the PCSCF checks the identity of the mobile home network to deduce the address of the ICSCF. Then, it initiates a secured session with the ICSCF based on TLS protocol. This process needs mutual certificates verification. Then, the Bridge Certification Authority (BCA) intervenes. The main purpose of the assumption on BCA is to facilitate the certificate verification process and to ensure inter-operability between the different operators.
Step 3. After the mutual authentication phase and the share of a symmetric session key, the PCSCF sends a message M_2 containing the information sent by the MS to the ICSCF.
Step 4. The ICSCF extracts the address of the HSS relative to MS from message M_2. Then, it initiates a secured session with this HSS based on TLS protocol. After that, the ICSCF retransmits to the HSS a message M_3 that includes the MS's information and the certificate of the current PCSCF.
Step 5. The HSS decrypts part E of message M_1 and verifies the signature of MS. Next, the HSS checks the private identity of MS and checks whether MS has the rights to accede the requested service. In the positive case, the HSS determines the address of the suitable SCSCF that is able to provide the service. On an other hand, the HSS verifies the validity of the PCSCF certificate and generates an identity certificate $Cert_{HP} = (K_{PPCSCF}, ID_{PCSCF}, \delta Cert_{HP})\_k_{pHSS}$ that will be transmitted to MS to verify the identity of the PCSCF. Variable $\delta Cert_{HP}$ defines the validity period of the certificate that we assume relatively short in order to avoid the verification of certificate revocation list at the MS level for this type of certificates. $Cert_{HP}$ is then sent to the ICSCF.
Step 6. On receipt, the ICSCF transmits certificate $Cert_{HP}$ and the public identity of MS to the PCSCF in message rM_5. rM_5 constitutes an implicit acknowledgment to the PCSCF, indicating that the identity of MS has been verified and that it should keep the connection until the accomplishment of the registration process. In the same time, the ICSCF establishes a secured session with the SCSCF based on the TLS protocol. Then, the ICSCF transmits message M_5 to the SCSCF which includes $\{E, Cert_{MS}, Cert_{HSS}, IMPU\}$.
Step 7. The PCSCF computes a symmetric session key K_{si} and signs it with private key k_{pPCSCF} and encrypts it with public key of MS K_{PMS}. Next, the PCSCF sends the $E' = [(K_{si})\_k_{pPCSCF}]\_K_{PMS}$, a challenge req_i and the certificate $Cert_{HP}$ delivered by the HSS to the MS in message rM_6.
Step 8. MS verifies the signature of the HSS on $Cert_{HP}$. It decrypts E' and verifies the signature of the PCSCF on the session key K_{si}. In the case of verification success, MS stores the session key to be used in its communications with the PCSCF. Furthermore, MS computes the response $res_i = (req_i)\_K_{si}$ and sends it to the PCSCF in the message rM'_6.
Step 9. When the SCSCF receives the message M_5, it extracts the address of the HSS, initiates a secured session with it based on the TLS protocol. Then, the SCSCF sends $E = [(IMPI)\_k_{pMS}]\_K_{PHSS}$ and the certificate $Cert_{MS}$ of the mobile.
Step 10. The HSS verifies the validity of mobile MS certificate and its private identity.
Step 11. The HSS extracts the authentication vectors relating MS. The structure and contents of these vectors are identical to those defined by the 3GPP in the IMS AKA (TS 33.203, 2002). The HSS extracts the different public identities of MS in order to give them to the SCSCF, to be used in the case where the same MS requests another access to a different service provided by the same SCSCF during the period of validity of the active registration. Then the vectors and the public identities are signed with the private key k_{pHSS} of the HSS and encrypted with the public key of the SCSCF. The HSS generates an attribute certificate in which it signs that the current SCSCF will offer the asked service to the MS $Cert_{HS} = (ID_{SCSCF}$,service,$\delta Cert_{HS})\_k_{pHSS}$, where $\delta Cert_{HS}$ is the validity period of the certificate, which is defined by the HSS in order to avoid the use of CRLs for MS. However, It should fulfill the following condition in order to guarantee service continuity:

$$\delta Cert_{HS} < \delta Cert_{HP}.$$

Step 12. The HSS sends the message $M_8 =\{ E" = (\{AV_i, \{IMPU\}\}_k_{pHSS})_K_{PSCSCF}, Cert_{HS}\}$to

the SCSCF. Moreover, it updates the mobile information (location, current request for registration, SCSCF concerned) and it is supposed to wait the end of the registration request to lunch the charging procedure.

Step 13. On the receipt of message M_8, the SCSCF decrypts E" and verifies the HSS signature. Then, the SCSCF selects a vector AV_i, extracts the values of $RAND_i$ and $AUTN_i$, which are sent with the $Cert_{HS}$ to the mobile MS.

Step 14. MS verifies the attribute certificate $Cert_{HS}$ and the freshness of the sequence number present in the AUTN$_i$. Next, it computes the response RES$_i$, the integrity key IK_i and the confidentiality key CK_i. Then, MS sends the response RES$_i$ to the SCSCF.

Step 15. The SCSCF verifies the correspondence of RES_i and the value $XRES_i$ present in the AV_i. On success, the SCSCF sends a flag M$_{15}$ to inform the HSS. It also sends a positive acknowledgment to the MS containing the period of time after which the mobile should proceed to a re-registration. This period is defined as $T - dt$ and is sent protected with the integrity key IK_i and the confidentiality key CK_i.

2.3 The Re-registration Protocol

When a mobile wants to extend an active registration, it proceeds as shown in Figure 2. For this, it attempts to perform the following steps:

- The MS sends its public identity and the last computed RES_i, for the current registration, protected with the integrity key IK_i to the SCSCF.
- When receiving this message, the SCSCF choses a new authentication vector AV_j and sends the corresponding $RAND_j$ and $AUTN_j$ to the MS.
- The MS computes the value RES_j and the new integrity and confidentiality keys (IK_j and CK_j). Then, it sends the RES_j to the SCSCF.
- After the verification process, the SCSCF starts the use of the new keys (IK_j and CK_j) and sends in an acknowledgment message to the MS the lifetime of the new registration.

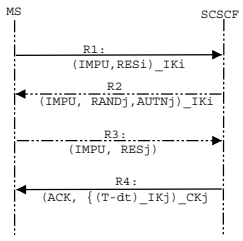


Figure 2: The re-registration procedure.

3 MANAGING SECURITY ASSOCIATION IN UMTS ENVIRONMENT

In 3GPP, security associations (SA) were only defined between mobiles and the PCSCF (TS 33.203, 2002). The setup of these SAs is done during the registration process. A SA includes the following five major attributes: (1) a uniquely defined identifier of the SA; (2) the destination and the source address or identifier; (3) the authentication, integrity and encryption algorithms; (4) the keys lengths; and (5) the finite SA lifetime.

We have found that this kind of SAs cannot be adopted as they are, since PCSCF has no longer the same responsibilities as those defined in (TS 33.203, 2002). Our aim in this section is to provide an adaptation of this paradigm to provide a good management process and better security of the exchanged elements.

3.1 Defining Security Associations

Since the integrity and confidentiality keys would not be sent to the PCSCF, new SAs should be defined between the MS and the SCSCF. Other SAs should take place between the MS and the PCSCF. These associations allow the definition of security agreements between the different communicating entities.

Managing SAs between MS and PCSCF. SAs ensure the establishment of the security mode allowing the access to the IP network. In fact, they would guarantee confidentiality and integrity of the exchanged data between the MS and the PCSCF. The associations are characterized by the following aspects:

- the identifier of the mobile is no longer its private identity $IMPI$. It will be replaced by the public identity ($IMPU$) and the IP Address of the mobile,
- the keys defined between the PCSCF and the MS are symmetric session keys (and no longer the integrity and the confidentiality keys).

The setup of SAs starts when the mobile sends its first request for registration. In fact, message M_1 includes the necessary information to negotiate security parameters between the MS and the PCSCF and to authenticate MS. Those information specify the identity of the mobile, its supported algorithms, and its identity certificate. Upon receiving message M_1, the PCSCF verifies the security mechanisms presented by the MS. Then, it waits for message rM_5 to authenticate MS. This message implies the result of the check of the validity of the identity certificate of the mobile as well as its private identity in the HSS. In addition, the PCSCF receives its temporary identity certificate delivered by the HSS and the current valid Certificate Revocation List (CRL). The two certificates allow the

PCSCF to set up the lifetime of the SA. Therefore, the lifetime of SA should be smaller than the validity period of both $Cert_{HP}$ and the CRL. Otherwise, SA would be invalid when one of the two certificates becomes invalid. In this case, a request for a certificate renewal or a new CRL should be sent.

After the determination of the SA lifetime, the PCSCF computes a symmetric key, signs it and encrypts it with the public key of the MS. Then it sends message rM_6 to the MS in which it defines the SAs. rM_6 would include the chosen security mechanisms, the certificate $Cert_{HP}$, the lifetime of the SA and the symmetric session key. To ensure the security of the SA, the two latter parameters should be sent signed and encrypted. In addition, the PCSCF sends a request req_i to MS. When receiving rM_6, the MS authenticates the PCSCF with the verification of $Cert_{HP}$. Then, it stores the parameters of the SA to be used on the following messages. Furthermore, MS computes the response res_i and sends it to the PCSCF to accomplish the SA setup procedure.

Managing SAs between MS and SCSCF. To provide an end to end security between the MS and the SCSCF, new SA should be established between these two entities. The definition of these SAs allows the protection of all kinds of access to the services independently of the access network. Even if there is a security weakness on the communication links, the SCSCF could verify, and no longer delegate, the integrity and confidentiality of the messages sent by MS. In this case, SAs are defined between the MS, which is defined by its public identity $IMPU$ and by its IP address, and the SCSCF defined by its IP address. The selection of the security mechanisms (e.g., authentication, integrity and confidentiality) to be used/declared in SAs is done during the registration process. In fact, the MS sends with message M_1 the lists of its supported security mechanisms, the index of its security association, its identity certificate and its root certificate. Based on M_1, the SCSCF determines the mechanisms that it would deploy. Then, it authenticates the MS through the verification of its private identity and its identity certificate in the relative HSS. After that, the SCSCF use the CRL and the validity period of the $Cert_{HS}$ to deduce the lifetime of the security association such as it will have the smallest value. Next, the SCSCF chooses an authentication vector AV_i and extracts $RAND_i$ and $AUTN_i$. These parameters would allow the MS to compute the integrity and the confidentiality keys IK_i and CK_i. Finally, the SCSCF sends the previous indicated parameters to the MS in message M_9. On the other side, the MS would verify the freshness of the message and the identity of the SCSCF. Then, it computes RES_i, IK_i and CK_i. Next, the response of the MS (using message M_{12}) will confirm the choices indicated in the SA.

To resume, one can note that the keys and the parameters defined in each security association are those existing in the authentication vectors delivered by the HSS to the SCSCF. Therefore, every authentication vector contributes to the definition of a security association. The lifetime of the SAs defined both in the MS and in the SCSCF are specified to be longer than the lifetime of the registration. Thus, the request for re-registration is protected by the security association yet established. After the definition of the two SAs, the next sub-section will consider the mechanisms of protection of the SA databases.

3.2 The Protection of the Security Associations

The security associations previously defined are stored in specific data bases (SADB). The protection of these databases needs in addition to a secured hard storage an enforcement of some appropriate protection. Thus, we propose the definition of two types of SADBs:

• the first contains the list of SAs established at a defined moment. It can include SAs established between the MS and the SCSCF:

SA#	Source Address	Destination Address	Encryption algo
Auth Algo	Integrity Algo	Ptr# rule#	

• the second contains the security policies defined between the different operators (e.g. between the SCSCF and the HSS).

Rule#	Source Address	Destination Address
Encryption algo	Authentication algo	Integrity Algo

The use of these databases ensures more protection of the SAs, since there is a continuous verification of the conformity of a security association to the rules defined between two different operators. Furthermore, this approach can offer security as a quality of service given to the subscribers according to the agreements defined between the HSS and the SCSCF.

4 SPUR ANALYSIS

In this section, the most important SPUR provisions are quoted. Furthermore, a comparison between the security mechanisms defined respectively in the 3GPP protocol and the SPUR is presented.

4.1 Security Provisions

SPUR scheme presents different security provisions. More precisely, one can state the following:

• SPUR guarantees the protection of the integrity and confidentiality of the transmitted data between the different entities of the IMS at the SIP layer. It ensures two types of security mechanisms. The first is based on the use of TLS between the nodes of the IMS including PCSCF, SCSCF, ICSCF and HSS. The second uses the different SAs established between the MS, PCSCF, and SCSCF. So that every kind of transaction between the different participants in a communication is highly protected.

• SPUR allows the MS to authenticate the SCSCF and the PCSCF in addition to the home network. This is ensured by the use of certificates delivered by the HSS to each node. The MS can verify each time the signature of the HSS on the identities of the two nodes. If the verification is successful, the MS is sure that the HSS had authenticated the PCSCF and the SCSCF.

• SPUR provides end-to-end security for the private data of the MS. In fact, the use of SAs between the MS and the SCSCF allows to the server as well as to the MS to be sure that the integrity and the confidentiality of the exchanged data are protected during the validity of the security association.

• SPUR is independent from the protocols used in the lower layers, since all the presented mechanisms are implemented in the SIP messages without a need for lower protocols layers.

• SPUR exploits (or integrates) the 3GPP registration procedure. In fact, we have not changed the authentication vectors defined by the GPP standards. However, we have added other mechanisms to enforce the security of the exchanged data.

4.2 Comparing SPUR to 3GPP Protocol

SPUR presents many enhancements comparing to the different propositions for UMTS registration. Siemens proposal has different drawbacks (S3-000689, 2000). First, there is no kind of authentication between the MS and the SCSCF. The MS only authenticates the HSS. This approach can induce different attacks such as masquerading and man in the middle attacks. Furthermore, the protection features used between the nodes of the IMS are based on the security mechanisms defined at the network layer. This means that the absence or the weakness of the security protocols implemented at the network layer could lead to an unprotected transmission of the private user data. This presents an important threat to the user security and does not respect the 3GPP requirements on SIP.

On an other hand, Ericsson proposal does not present a practical solution (S3z000010, 2000). In fact, it adds significant loads to the HSS, which must insure authentication and re-authentication each time a mobile accesses the IMS. Also, the performance of the HSS would decrease when sending the challenge and waiting for a response. Another drawback is related to the complexity of the re-authentication procedure, which is invoked by the HSS and induces the retransmission of the new integrity and confidentiality keys to the SCSCF and PCSCF.

3GPP has defined an other registration protocol that overcomes many drawbacks defined in previous proposals. It insures the authentication of the MS to the SCSCF, which receives the authentication vectors from the HSS. However, the PCSCF terminates the integrity and confidentiality protection using the appropriate keys defined by the authentication vectors. 3GPP protocol has also some drawbacks. First, the MS does not identify the SCSCF. Second, the private identity of the MS is clearly transmitted in the SIP layer for each authentication or re-authentication procedure. Third, the integrity and confidentiality keys of the user are transmitted from the SCSCF to the PCSCF using the network layer security mechanisms.

The following table summarizes the security enhancements provided by the SPUR scheme in comparison with the 3GPP registration procedure. The table shows in particular that SPUR provides at least six additional security services including SIP authentication and confidentiality.

Table 1: Comparison of the security provisions of the 3GPP and the SPUR.

Criteria	3GPP	SPUR
SIP Authentication	1	1
SIP Confidentiality	0	1
SIP Integrity	0	1
IMPI Confidentiality	0	1
Establishment of the IK and the CK	1	1
Key Freshness	1	1
Serving Network Authentication	0	1
Certification use	0	1
Key Session Definition	0	1

5 SPUR SIMULATIONS

5.1 Simulation Environment

In this subsection, the impact of the addition of new processing in the IMS nodes is studied. Particularly,

we will focus on the influence of the changing size of the signaling messages respectively on the error probability and on the data flows exchanged between the different nodes.

The simulation model we use is based on the studies defined in(Kist and Harris, 2002; Handlay et al., 1999). We have applied SPUR on four nodes which represents the MS, the PCSCF, the ICSCF and the SCSCF. The arrival process to IMS nodes is the Poisson process with a mean arrival rate of 1 session per second. IMS nodes have also a negative exponential service times with means of 20ms. We consider that the original size of a SIP message is varying between 170 bytes and 500 bytes(Rosenberg et al., 2003)(Sweeny et al., 2003). Also, cases where the additional size increased per each node is variable between 50 bytes and 200 bytes according to the type of the processing are studied. The first simulation considers the error probability defined for the transmitted messages depending on the Error Bit Rate (BER) and the size of messages. The error probability of a message is defined using the binomial distribution $P_E(M) = \sum_{k=1}^{8M} BER^k(1-BER)^{8M-k}$ where M is the size in bytes and k is the number of corrupted bits. The second study determines the additional flow defined on the links between the different nodes. Let's define n as the number of the links and l_i as the link on which the flow is calculated. The flow defined in one direction on link *li* is $F(l_i) = M(l_i)(1 + \sum_{m=l_i}^{n} \frac{P_E(m)}{\prod_{j=l_i}^{m}(1-P_E(j))})$.

5.2 Analysis of the Numerical Results

We present in this subsection the results obtained by the execution of the previous model.

• **Error probability**

The following figures present the impact of the augmentation of the size of the messages on the error probability. Two cases are studied: the first considers the high error bit rates (shown in Figure 3) while the second presents the variation of the error probability for low values of BER (Figure 4). We notice that for low values of BER, the error probability is almost linear. Hence, to have reasonable error probabilities (i.e., less than 10^{-3}) the values of BER should be chosen lower than 10^{-6}. However, if BER$\geq 10^{-5}$, the error probability takes important values and grows exponentially toward the maximum probability even for small sizes.

Thus, we can conclude that we should choose the constraint (BER <10^{-5}) to have an acceptable error probability, and so have less message retransmission between the nodes.

• **Additional Flow**

The changing size of the transmitted messages induces additional flow between the network nodes, especially when the BER is not the same on the different links. In the following figures, three cases are considered: The first case considers the BER on the links between the nodes has high values (Figure 5). The second considers the same BER on all links (Figure 6), and the last addresses the case where BER has small values between links (Figure 7)e

The first figure shows the case where the radio link has a great value of BER. We can notice here that the additional flow on the link can reach 180% of the initial flow size which is unacceptable on the radio link since it adds unacceptable amounts of interference.

In the second case, we study the case where the BER is the same on all the links. This situation is not usually true since the radio link presents always the highest BER. Nevertheless, we notice that for a BER = 10^{-6}, the additional flow overheat has a maximum value of 8% .

The last studied case considers low values of BERs. We notice in this case that the additional flow does not take important values. It can be induced that the use of SPUR with low BERs on the links between the different nodes does not add high flows. The previous results demonstrate that the use of SPUR does not introduce large loads if the BER values are well chosen.

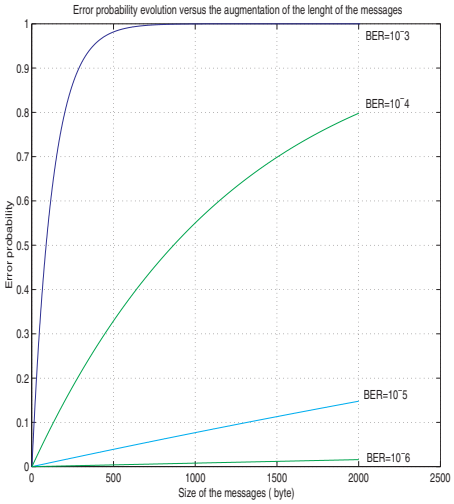


Figure 3: Error Probability in the case of high BER.

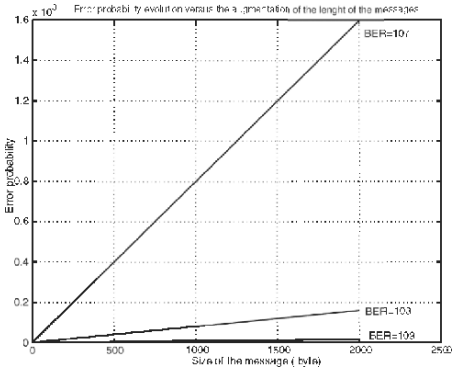


Figure 4: Error Probability in the case of low BER.

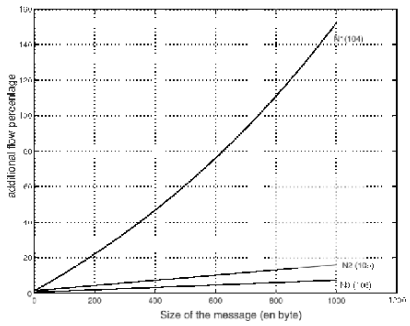


Figure 5: Additional ßow for high BER.

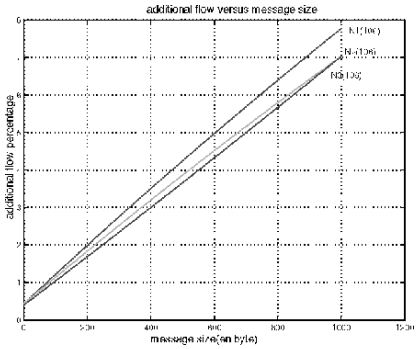


Figure 6: Additional ßow for equal BER.

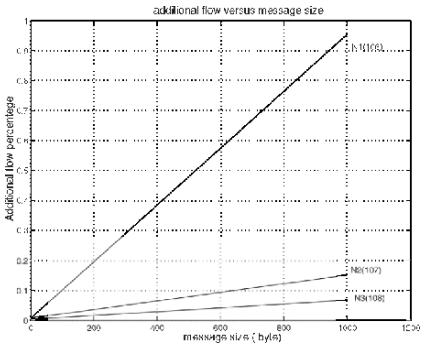


Figure 7: Additional ßow for low BER.

6 CONCLUSION

In this paper, we have presented a secure registration protocol for UMTS all IP networks. This protocol has added new security measures that provide mutual authentication, integrity and confidentiality between all entities involved in a communication process. In addition, it provides an end-to-end security service for the mobile privacy.

SPUR scheme is an extensible protocol that can define a comprehensive platform to integrate next generation networks (NGN), assuming that they are based on SIP-like protocols. The integration would assume that a bridge architecture of certification is made available in a way that any certificate provided can be checked efficiently.

REFERENCES

Kaaranen, H., Ahtiainen, A., Laitinen, L., Naghian, S., Niemi, V., (2001). *UMTS Networks : Architecture, Mobility and Services,* Weily, England,

TS 23.228: IP Multimedia subsystem Stage 2. Retrieved March 3, 2003, from http ://www.3gpp.org

TS 22.228: Service Requirements for the IP Multimedia Core Network. Retrieved June 6, 2002, from http ://www.3gpp.org

TS 33.900: A Guide to 3rd Generation Security. Retrieved January 1, 2000, from http ://www.3gpp.org

TS 33.120: UMTS Security principals and objectives. Retrieved May 5, 2000, from http ://www.3gpp.org

TS 21.133: *Threats and attacks in UMTS*, Retrieved December 12, 2001, from http ://www.3gpp.org

TS 33.203: Access Security for IP-based services. Retrieved March 3, 2002, from http ://www.3gpp.org

TS 24.229 IP Multimedia Call Control Protocol based on SIP and SDP, V5.4.0 Retrieved March 3, 2002, from http ://www.3gpp.org

Rosenberg, J., Schulzrinne, H., Camarillo, A., Johnston, G., Peterson, R., Sparks J., Handley, M., . Schooler, E, (2002). *RFC 3261: SIP. Session InitiationProtocol* Retrieved August 8, 2003, from IETF web site: http ://www.ietf.org/rfc/rfc3261.txt

3GPP TSG SA WG3 Security, S3-000689 : Access security for IP-based services. Retrieved November 11, 2000, from *www.3gpp.org/ftp/tsg_sa/WG3_Security/ 2001_meetings/TSGS3_17_Gothenberg/Docs/PDF*

3GPP TSG SA WG3 Security, S3z000010. 2000. Authentication and protection mechanisms for IM CN SS; Retrieved November 11, 2000, from *www.3gpp.org/ftp/tsg_sa/WG3_Security/ 2001_meetings/TSGS3_17_Gothenberg/Docs/PDF*

TS 33.102: *3G Security, Security Architecture*, Retrieved September 9, 2000, from http ://www.3gpp.org

Hastings N.E., Polk W. T. (2000), *Bridge Certification Authorities : Connecting B2B Public Key Infrastructures*. Retrived from csrc.nist.gov/pki/documents/B2B-article.pdf

TS 23.060: General Packet Radio Service; Service Description. Retrieved March 3, 2002, from http ://www.3gpp.org

Kist, A., and Harris, R.J., (2002). A Simple Model for Calculating SIP Signaling Flows in 3GPP IP Multimedia Subsystems, *Lecture Notes in Computer Science,* 924-935

Handlay, M., Schulzrinne, H., Schooler, E., and Rosenburg, J.D., (1999). *RFC 2543*: SIP: Session Initiation Protocol . Retrieved March 3, 2003, from IETF web site : http ://www.ietf.org/rfc/rfc3261.txt

Sweeny, J., Kenneally, V., Pesch, D., Purcell, G, (2003). *Efficient SIP based Presence and IM services with SIP message Compression in IST OPIUM.* Retrieved September 9, 2003, from http ://www.ist-opium.org/

PROVIDING QOS IN 3G-WLAN ENVIRONMENT WITH RSVP AND DIFFSERV

Eero Wallenius
Nokia Networks/ OS, Hatanpäänvaltatie 30, FIN-33100 TAMPERE
Tel:+358 40 5004055, Fax:+358 7180 74385
eero.wallenius@nokia.com

Timo Hämäläinen, Timo Nihtilä and Jyrki Joutsensalo
University of Jyväskylä, Department of Mathematical Information Technology, Finland
{timoh, nihti, jyrkij}@cc.jyu.fi

Keywords: 3G, WLAN, QoS, 3G-WLAN Interworking.

Abstract: Here we present the end-to-end QoS mechanism in 3G-multiaccess network environment. As multi-access wireless WLAN and wired xDSL wideband multi-access technologies has emerge and become more popular a need for interoperability with different technologies and domains has become necessity. There is also a need for end-to-end QoS management. We show a scenario where the UE-GGSN connection is covered by RSVP and RAN network part uses partial over dimensioning and real-time controlled ATM queuing. DiffServ covers WLAN-Core QoS and radio interface between WLAN AP and WLAN UE uses IEEE's 802.11e. Our interest is to find out how well 3G traffic classes can survive in different traffic conditions in the end-to-end case.

1 INTRODUCTION

With the evolution of QoS-capable 3G wireless networks, the wireless community has been increasingly looking for a framework that can provide effective network-independent end-to-end QoS control. One bigger problem arises with this kind of diverse networks. It is the dissimilarity of traffic characteristics and QoS management methods. Problem with WLAN networks is the high error rate probability. 802.11e standard has been applied trying to correct the situation by enabling the use maximum of eight separate priority queues for prioritizing higher priority traffic compared to other traffic [802.11e]. QoS supported WLAN uses the Enhanced Distributed Coordination Function (EDCF). It is the basis for the Hybrid Coordination Function (HCF) [802.11e]. Our research is also related to 3GPP WLAN interworking standardization [TS23.234].

RSVP has been used in domains where there is no direct radio interface. In the RAN case we have assumed that the radio interface between BTS and UE in RAN will be handled similarly to WLAN but with different methods defined by 3GPP standardization. As RAN is based on ATM the basic assumption has been that the RAN correctly dimensioned to carry al traffic coming from and going to UE direction so by default RAN QoS is out of scope of scenarios in this paper.

This research is part of the 3G–WLAN Interworking research program made during years 2002-2004 [Wallenius E., Hämäläinen T., Nihtilä T., Luostarinen K.] and [Hämäläinen T., Wallenius E., Nihtilä T., Luostarinen K.]

2 MAPPING QS ATTRIBUTES TO CROSS DOMAIN INTERFACES

3GPP has defined four traffic (QoS) classes and three subclasses (Interactive THP, Traffic Handling Priority) that can have their own QoS attributes [TS 23.107]. All traffic in the 3G network will be handled according to the operator and service's requirements at the each of these traffic classes. The main QoS method to be used at the core network is

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 211–217.

supposed to be DiffServ [TS22.934]. Addition to that 3GPP has defined RSVP as an additional UE originated QoS method [TS23.917] in Rel6 between UE-SGSN and GGSN. It can be used at the situations where scalability problems will not arise (small networks). 3G traffic classes are: Conversational class for voice and real-time multimedia messaging. Streaming class for streaming applications (Video On Demand (VOD) etc.), Interactive class for interactive applications (eCommerce, WEB-browsing, etc.), Background class for background applications such as email and FTP. QoS values for each traffic classes are defined in [TS23.107]. In DiffServ domain four priority queues can be implemented for the each 3G traffic classes. The three THPs (Traffic Handling Priority) are also available for Interactive class to further sub-classify Interactive class traffic by inserting it to three separate queues. 3G to DiffServ mapping process can be policy based controlled and the mapping can be indicated at the IP level by the DSCP (DiffServ Code Point) inserted to the TOS field by DS classifier/marker mechanism or by the actual application that generates the control plane traffic. Table 1 shows the PHB actions with DSCP mappings.

The nature of RSVP functionality differs significantly from DiffServ. RSVP uses end-to-end signaling enabling a single UE to reserve end-to-end transport capacity from the network or RSVP can be used by Bandwidth Broker and COPS-PR protocol to set appropriate traffic filters to routing nodes to achieve similar capacity reservation than by UE signaling.

3 SIMULATION ENVIRONMENT AND PARAMETERIZATION

The goal is to study what are throughputs, delays and dropping rates in RSVP and DiffServ cases. Simulation environment in Figure 1 consists of 18 Access points which each connected to UEs with different traffic priorities. Six core network nodes build up a ring and each of them has three access points. WLAN stations send data at the rate of 2.5Mb/s. Stations no. 1 and no. 3 generate CBR traffic and stations no. 2 and no. 4 send VBR traffic. The stations start sending at time interval 3-4.5 seconds randomly. Simulation time is 40 seconds and the used packet size is 1000 bytes for all stations..

Table 1: RSVP parameterization.

3GPP Traffic class	Bandwidth Mb/s	Bucket size bytes
Conversational	3.0	3000
Streaming	2.5	2000
Interactive (3 THPs)	2.0	1500
Background	2.0	1500

UEs for AP 1-9 are sending and 10-18 receiving. Available bandwidth within the core network was 8 Mb/s.

In the core network all wired capacity was reserved for RSVP use and best effort queue size was 5000 bytes in every node.

We used traffic parameterization shown in Table 1.

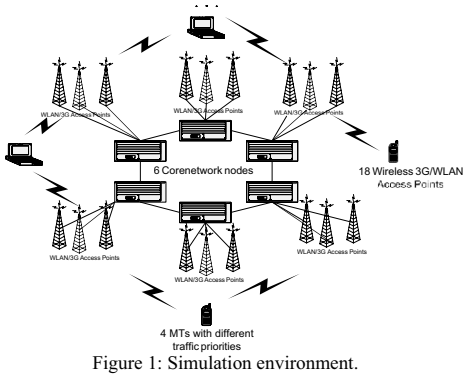


Figure 1: Simulation environment.

As link capacity is small compared to number of reservations some of the reservations does not success and traffic related to them goes in the network as best effort traffic. RSVP uses WFQ queuing. DiffServ uses Token Bucket Polices and its parameterization is presented in Table 2.

Table 2: DiffServ token bucket parameterization.

3GPP Traffic class	CIR Mb/s	Bucket size bytes
Conversational	3.0	3000
Streaming	2.5	2000
Interactive (3 THPs)	2.0	1500
Background	2.0	1500

.

DiffServ uses RED queuing in drop tail mode. In-profile packet queue lengths are 30 packets for each class and out-of-profile packet queues are 60 packets long.

We used four priority levels in both scenarios. EDCF parameters of different Traffic Classes are shown in the following Table 3.

Table 3: EDCF parameters.

3GPP Traffic class	Conv.	Stream	Interact.	Backgr.
CWMin	7	10	15	127
AIFS (CWOffset)	2	4	7	15
CWMax	7	31	255	1023

To emulate the process of packet transmission errors we extended the simulator by implementing a two-state Markov model in the air interface. In our error scenario, the channel switches between a ”good state” and a “bad state”, G and B respectively: Packets are transmitted correctly when the channel is in state G, and errors occur when the channel is in state B. When the channel is in state G, it can either remain in this state, with probability ω_1 or make the transition to state B, with probability 1-ω_1. Likewise, if the channel is in state B, it remains in this state with probability ω_2 and changes state with probability 1-ω_2.

Table 4: Transition probabilities for 2-state MMPP.

Error rate	ω_1	ω_2
0%	0	1
20%	0.16	0.63

All test were done with network simulator NS-2 with IEEE 802.11 EDCF functionality implemented by Project-INRIA [Ni Qiang]. We ran several different error rate scenarios but we find 0 and 20% error rates most illustrative.

3.1 Scenario 1: RSVP Case

3.1.1 RSVP Throughputs

As can be seen in Figure 2 Interactive class has higher throughput than Streaming class.

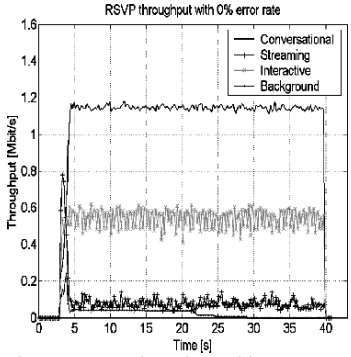


Figure 2: RSVP throughput with 0% error rate.

This is caused by the random nature of reservation signalling.

The reservation probabilities are shown in Table 5.

In case that there is already 6Mb/s reservation for two Conversational class flows only Interactive and Background classes can reserve the rest of the bandwidth.

Other traffic characteristics follow very well expectations on throughput delay.

Table 5: Reservation probabilities.

Mb/s	Conv.	Stream	Interact.	Backgr.
8	0.25	0.25	0.25	0.25
6	0.25	0.25	0.25	0.25
5.5	0.25	0.25	0.25	0.25
5	0.25	0.25	0.25	0.25
4	0.25	0.25	0.25	0.25
3.5	0.25	0.25	0.25	0.25
3	0.25	0.25	0.25	0.25
2.5	0	0.33	0.33	0.33
2	0	0	0.5	0.5
Average	0.194	0.231	0.287	0.287

Throughput is best and delay follows the throughput being higher than in other classes due to the high throughput.

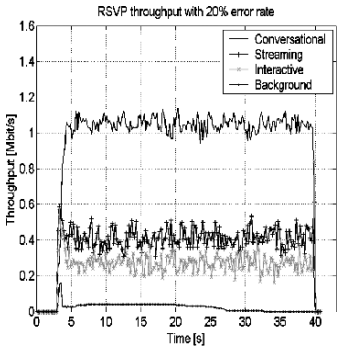


Figure 3: RSVP throughput with 20 % error rate.

Also can be seen in Figure 2 and Figure 1 that the traffic flows are smoother in lower error environment.

Average throughputs on each traffic class also follow well our expectations. Throughputs are in preferable order, Conversational highest and background lowest.

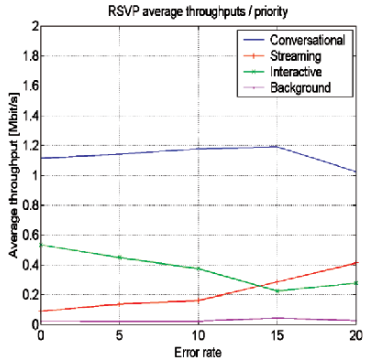


Figure 4: RSVP average throughputs / priority.

Figure 4 shows also slight rise of throughput in Interactive class and corresponding declining in Background class for higher error rates. This can be caused by differences in reservation success between classes.

3.1.2 RSVP Delays

Delay behaviour is similar as throughput. All aggregate flows, traffic classes, are in correct order and delay is adequate low (<0.5 ms) in both Conversational and Streaming class for their 3G usages. Also Interactive and Background classes are far below their worst-case scenario values, several seconds. See Figure 5 and Figure 6.

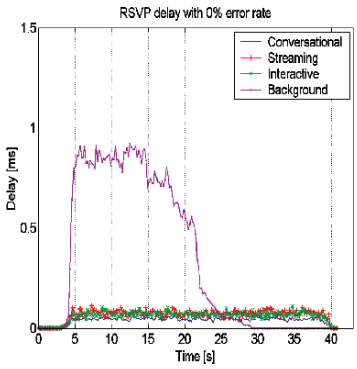


Figure 5: RSVP delay with 0% error rate.

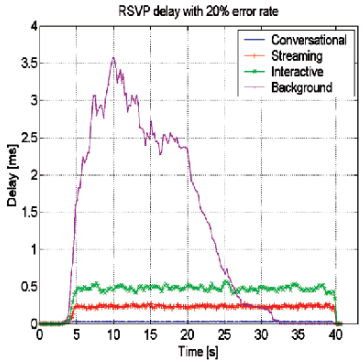


Figure 6: RSVP delay with 20% error rate.

3.1.3 RSVP Packet Droping

RSVP packet dropping follows the throughput being higher in higher throughput classes as expected. In this case a better describer for packet dropping would probably be percentage value, which would turn the order of curves into opposite order.

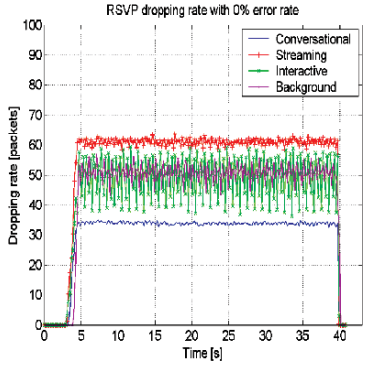


Figure 7: RSVP dropping rate with 0% error rate.

Dropping rate is very stable when the dropping rate is 0% Figure 7 but becomes unstable and rising with error rate 20%.

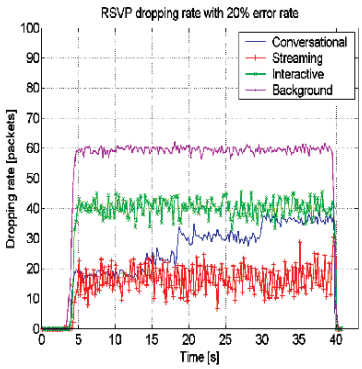


Figure 8: RSVP dropping rate with 20% error rate.

3.2 Scenario 2: DiffServ case

3.2.1 DiffServ Throughputs

DiffServ show different kinds of throughput results than RSVP. Conversational traffic is dominant and other traffic classes are very close to nil. The obvious difference is that RSVP has much better control over lower priority flows and therefore it would be a better solution for Interworking QoS control purposes.

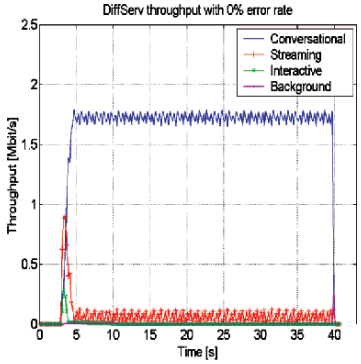


Figure 9: DiffServ throughput with 0% error rate.

As can be seen form Figure 9 traffic with priorities 3 and 4 disappears within 10 seconds after beginning of the test. This means also that the delay

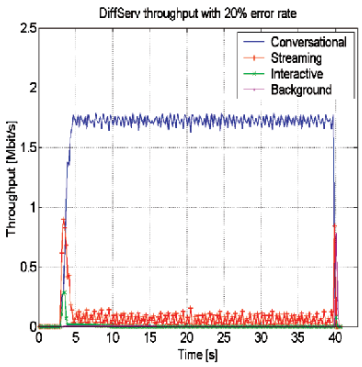


Figure 10: DiffServ throughput with 20% error rate.

for priorities 3 and 4 becomes 0 (zero), as there is no traffic in priority classes 3 and 4 as shown in next chapter.

Difference between throughputs with 0% and 20% error rate is significantly low.

Figure 11 shows that the average throughputs of the classes are the same between different error rates. This indicates that throughput behavior is very stable when using DiffServ in opposite to RSVP, which causes large variations in class throughputs between error rates.

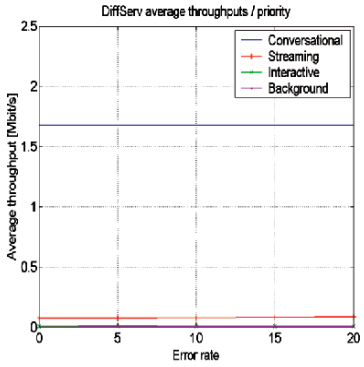


Figure 11. DiffServ average throughputs / priority.

Still, this stable behavior is achieved in cost of lower priority class throughputs, which are close to zero.

3.2.2 DiffServ Delays

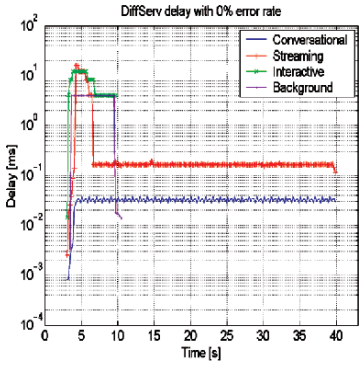


Figure 12: DiffServ delay with 0% error rate.

As presented in Figure 12 the delay for flows with priorities 3 and 4 become zero (vanishing from logarithmic scale). This actually means that after a few seconds after stations have started to send flows with priorities 3 and 4 are not reaching their target receiver node but are totally dropped during transmission. Similar effect occurs with 20% error rate in Figure 13.

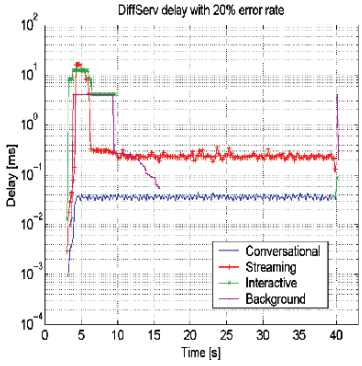


Figure 13: DiffServ delay with 20% error rate.

3.2.3 DiffServ Packet Droping

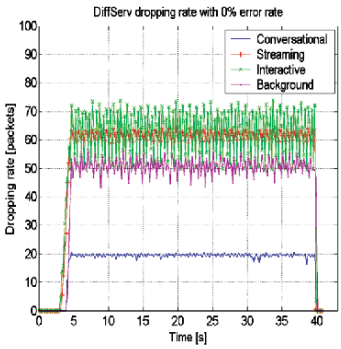


Figure 14: DiffServ dropping rate with 0% error rate.

As shown in Figure 14 dropping rates are located as could be predicted according to their priorities.

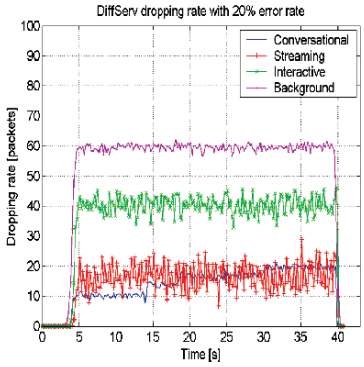


Figure 15: DiffServ dropping rates with 20 % error rate.

Naturally as presented in Figure 15 increased error rate increases dropping rate accordingly. High air interface error rate affects the dropping rates, so that there seems to be lower dropping rate in 20% error rate scenario. As the air interface corrupts packets, fewer of them reach the wired network. Hence, there is smaller probability of congestion in the wired network.

3.3 Test Conclusions and Recommendations

3.3.1 Combined Throughputs

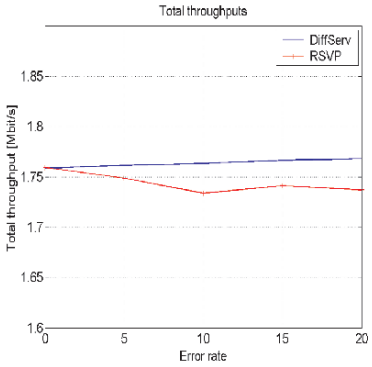


Figure 16: Comparison of total throughputs with RSVP and DiffServ.

Figure 16 shows that the throughput in DiffServ case is slightly better than in RSVP case. That is expected due to the resource reservation nature of RSVP. In DiffServ case all traffic classes can have unlimited number of flows compared to RSVP's bandwidth limiting functionality and access control. The difference between these techniques is almost negligible due to the fact that both RSVP and DiffServ achieve the maximum capacity of the network. This is due to the amount of traffic in the network: the flows are sending traffic so intensively that there is always a demand of bandwidth for best effort traffic and hence the network is never idle.

4 CONCLUSIONS AND FUTURE WORK

4.1 Achievements

In this paper we provided architecture for end-to-end QoS control in a wired-wireless environment with effective QoS translation. We used DiffServ and RSVP in the core network and 3G/WLAN and 802.11e at the wireless part of the tests.

Results show clearly that RSVP can keep delays smaller than in the DiffServ case. Paper also shows that the best and most suitable combination of QoS control would be RSVP-802.11e hybrid. Suitability materializes especially in the control of lower priority flows enabling them more and more controllable bandwidth with lower and controllable delay.

4.2 Future Studies

Next we will expand our simulations to cover a real operating size network and study how the operating parameters can be tuned e.g. by using dynamic policy based management.

Also further development of 3G interworking with other access methods is gaining increasingly importance and to achieve solid and robust Interworking QoS is the next top research challenges for the future.

5 REFERENCES

TS22.934, *3GPP Technical Specification,* "Feasibility study on .

3GPP system to Wireless Local Area Network (WLAN) interworking", Release 6.0.

TS23.107, *3GPP Technical Specification,* "QoS Concept and Architecture", Release 5.7.

TS23.917, *3GPP Technical Specification,* "Dynamic policy control enhancements for End to end Quality of Service (QoS)" V1.2.0.

TS23.234, *3GPP Technical Specification*, "3GPP system to Wireles Local Area Network (WLAN) interworking; System description", Release 6.0.

802.11e, IEEE WG, Draft "Supplement to Standard for Telecommunications and Information Exchange Between Systems - LAN/MAN Specific Requirements" - *Part 11: wireless Medium Access Control (MAC) and physical layer (PHY) specifications:* Medium Access Control (MAC) Enhancements for Quality of Service (QoS), IEEE 802.11e/D2.0, Nov. 2001.

TS 23.234, *3GPP Technical Specification* "3GPP System to Wireless Local Area Network (WLAN) Interworking: System Description", v1.0.0, September 2002. (Release 6)

Ni Qiang, http://www-sop.inria.fr/planete/qni/Research. html

Hämäläinen T., E. Wallenius E., Nihtilä T., Luostarinen K., "Providing QoS at the Integrated WLAN and 3G Environments", *14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC2003),* Oct. 2003.

Hämäläinen T., Wallenius E., Nihtilä T., Luostarinen K., "End-to-End QoS Issues at the Integrated WLAN and 3G Environments", *The 9th IEEE Asia-Pacific Conference on Communications (APCC),* Sept. 2003.

Wallenius E., Hämäläinen T., Nihtilä T., Luostarinen K., "3G Interworking with WLAN QoS 802.11e", *IEEE Semiannual Vehicular Technology Conference, (VTC'2003Fall)*, Oct. 2003.

FAST MOBILE IPV6 APPROACH FOR WIRELESS LAN BASED NETWORKS

Link-Layer Triggering Support for IEEE 802.11

Norbert Jordan and Alexander Poropatich
Institute of Broadband Communications, Vienna University of Technology
Favoritenstrasse 9/388, A-1040 Vienna, Austria
Email: *norbert.jordan@ieee.org, alexander.poropatich@tuwien.ac.at*

Keywords: Mobile IPv6, Fast MIPv6, Wireless LAN, Link-Layer Triggering, Fast IEEE 802.11 Handover.

Abstract: The standard Mobile IPv6 specification provides comprehensive mobility management for the IPv6 protocol. During the handover there is a period in which the mobile node is unable to send or receive packets due to link-layer switching and IPv6 protocol layer operations. This overall handoff latency resulting from baseline MIPv6 procedures, namely movement detection, new care-of address configuration, and binding updates with peer entities, is often unacceptable for any kind of real-time service (video-conferencing, voice-over-IP,…). A new fast handover approach, based on Fast Handovers for Mobile IPv6, is proposed in this paper, which will support seamless movement in between IPv6 domains using a IEEE 802.11 network infrastructure. A new low latency handoff method for IEEE 802.11 will be proposed, where access point beacons are utilized for carrying IPv6 prefix information without altering the Mobile IP or IEEE 802.11 specifications. A WLAN service will continuously monitor the radio signal quality of the attached access point and, if necessary, will switch to another access point in range. This feature and the elimination of firmware-based active scanning during link-layer handovers have the flavor effect of reducing the overall link-layer handoff delay to about 10%. We will further introduce our wireless testbed infrastructure for evaluation of the proposed approach. Performance evaluation is used to verify the effectiveness of our implementation and an extensive simulative comparison is used for scalability analyses.

1 INTRODUCTION

Owing to the assistance of Mobile IPv6 (Johnson, 2004), a mobile node can effectively maintain its IP-layer connectivity to the Internet when it changes its point-of-attachment somewhere in the world. During the accomplishment of the handover, the mobile node is unable to send or receive IPv6 packets because of its L2 and also L3 handover operations. This high handover latency is unacceptable to real-time applications or delay sensitive traffic. Each time a mobile client moves, it is necessary to perform movement detection by discovering (sending router solicitation) its current point of attachment. In Mobile IPv6 (Johnson, 2004), the movement detection algorithm relies on the periodic sending of router advertisements in order to enable the mobile node to determine its current location. The only way to improve the detection performance is to broadcast router advertisements at a faster rate, which may result in a poor link utilization. For that reason the fast handover protocol (Koodli, 2004) is designed to achieve a seamless handoff when mobile nodes move from one domain to another.

In a mobile-initiated and anticipated fast-handover scenario described in (Koodli, 2004), the mobile node first sends a Router Solicitation for Proxy (RtSolPr) message to the current access router containing any Access Point specific identifiers. The current Access Router replies with a Proxy Router Advertisement (PrRtrAdv) message, which may contain a subnet-specific information tuple [AP-ID, AR-MAC, AR-IP]. This message exchange allows a mobile node to obtain the new Access Router's prefix information, which is needed to perform an "anticipative" configuration of the new IPv6 address on the new subnet. Figure 1 presents a general mobile-initiated "predictive" fast handover scenario.

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 219–225.

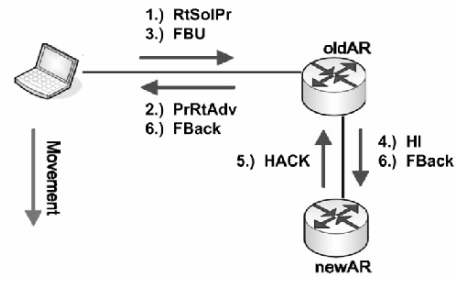


Figure 1: Reference Scenario for FMIPv6 Handover.

With the information provided in the PrRtAdv message, the MN formulates a prospective new CoA and sends a Fast Binding Update (FBU) message. The purpose of FBU is to authorize the old AR to bind the current Care-of address (CoA) to new CoA, so that arriving packets can be tunneled to the new location. Depending on whether an FBack (Fast Binding Acknowledgement) is received prior to the Mobile Node's movement or not, the prospective address can be used immediately after attaching to the new subnet link. In case it moves without receiving an FBack, the MN can still start using the new CoA after announcing its attachment through a Fast Neighbor Advertisement (FNA) message (see Figure 2)

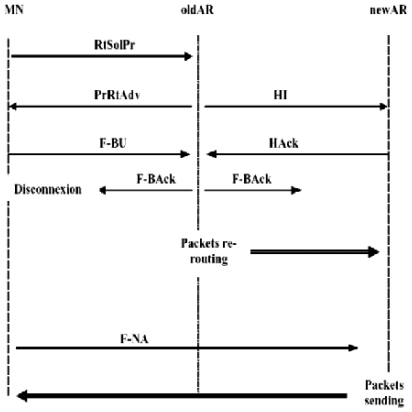


Figure 2: Message Flow for Mobile-Initiated HO.

However, the above protocol assumes that the L2 protocol is capable of delivering the L2 identifier of the new access point to the mobile node. More important, to initiate a seamless handover, is the fact that the current AR must be capable of mapping this new L2 identifier into the IP address of the target AR. We will show that all these requirements for Fast MIPv6 can be fulfilled in our implementation without any modifications to the IEEE 802.11 standards.

2 FAST HANDOVER FOR IEEE 802.11

The growing popularity of IEEE 802.11 (IEEE, 1999) has made "wireless" LAN a potential candidate technology for providing high speed reliable wireless access services. In addition by supporting Mobile IP, wireless LAN can meet demands for expanded wireless access coverage while maintaining continuous connectivity from one domain into another. In order to be able to accomplish a fast handover on Layer 3 it is necessary to implement a triggered information indicated by the underlying link-layer driver.

2.1 Link-Layer Triggering

In order to achieve an efficient interworking between Fast Mobile IPv6 and IEEE 802.11, it is necessary that the link-layer initiates the handover. The mobile node normally does this by sending a proxy router solicitation at the IP layer. This action is triggered by the underlying link layer in the mobile node, which must be aware that a handover is about to take place. This is the only possible way since from the IEEE 802.11 link-layer's point of view the mobile node is the only entity which is aware, that the host is about to attach to a new AP. In our implementation there is a tool running at the mobile node which continuously monitors the signal strength of the attached AP. In case the receiving power-level falls below a pre-defined value, the tool reacts by collecting information of all APs in range. So the tool is able to anticipate the best destination for the handover. At the same time of preparing the link-layer handover to the most qualified AP, the client-tool will send a trigger message to the fast-handover module. The next step that follows is the proposed FMIPv6 approach explained in Section 3.

2.2 Enhanced WLAN Handover

Even if the Fast Mobile IP approach is implemented properly, there are still delay issues to solve during the link-layer handover. Since Mobile IP and link-layer handover should go hand-in-hand, there is still an unsolved problem with the Layer 2 handoff-latency when the mobile node moves from one AP to another. There exists a definite period of time in which the mobile node is unreachable due to the

layer 2 movement (i.e. re-synchronization with the new AP). It has to be remarked that the exact amount of time varies, depending on the deployed WLAN technology. Some measurements (Velayos, 2003) (Velayos, 2004) (Mishra, 2002) for IEEE 802.11b show that this time period can vary from 200 to 1500 ms, depending on the type of vendor equipment.

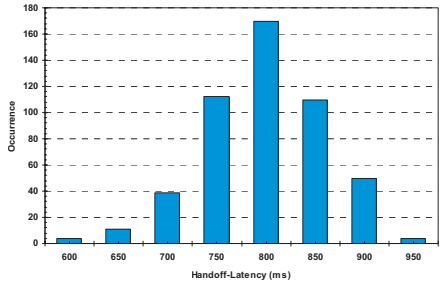


Figure 3: IEEE 802.11b HO Latency without Optimization.

The main problem during the handover is the fact that stations have to detect the lack of radio connectivity based on unsuccessful frame transmissions. The difficulty is to determine the reason for the failure among collision, radio signal fading or the station being out of range. In our implementation the signal strength is monitored continuously. In case that the signaling level drops below a predefined threshold, the tool automatically tries to handoff to an AP in range, which provides a much better connectivity. So the long phase of detection can be saved and the handover is carried out much faster. This WLAN handover-tool takes advantage of the information provided by the physical layer and completely skips the detection phase. Stations equipped with our tool start the search phase when the quality of the radio-signal falls below a pre-defined threshold. Therefore, the search always starts before any frame has been lost. This has the favorable effect that the overall handover-time can be reduced to about 350ms, as demonstrated in (Jordan, 2003). Another issue of WLAN is the active-scan process, which is often enforced with each AP-handover. Preventing active-scanning, additionally helps to reduce the link-layer latency to about 60 to 100 ms (depending on vendor hardware).

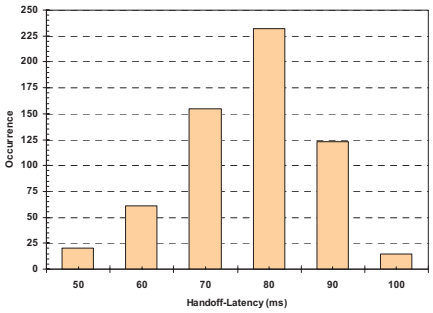


Figure 4: Optimized IEEE 802.11b Handover.

These initial improvements will enable wireless networks to carry real-time applications along the infrastructure.

3 PROPOSED FAST HANDOVER APPROACH

As already stated in Section 2 our implementation helps the Mobile Node to detect if the current link is degrading and therefore starts searching for a new AP with improved link-quality. To do this, the mobile node scans all possible frequencies (specified by the IEEE 802.11b standard) [10] and compares the received signal with the one currently received. If the mobile node finds a better signal it can switch to the new AP. But the mobile node's link layer implementation does not know whether this AP is attached to a new AR. The link layer only knows about link layer addresses and the AP's SSID (Service Set Identifier) string. However, if the AP name/link layer address (which identifies an AP) is known, the mobile node's IP-layer implementation can request that the current AR should provide the prefix/router address, which the new AP is attached to. This idea assumes that an AR is configured with a table containing its own and the neighboring APs link-layer addresses and their corresponding AR.

In our implementation we configure each access point involved with a special SSID string (e.g.: SSID = "2001:200:8:72AB:1434::1/64") which further implicitly presents all information about the prefix of the attached AR. Whenever the mobile node anticipates a handoff, the handover-tool exactly knows the prefix of the new AR the AP is attached to. In that way the mobile node performs "anticipative" configuration of the new IP address on the new subnet using the router prefix information carried in the beacon message of the new AP. If more than one

destination access point is in range, the mobile node could prefer to carry out a movement to an AP within the same subnet. Thus only a link-layer handover would be performed, which further improves the handoff-latency in this special case. In all other cases the mobile node will perform the configuration of a new IP address and continues with the Fast Mobile IPv6 handover until the mobile node arrives at the new AR (NAR).

4 IMPLEMENTATION OVERVIEW

To make a serious network evaluation in the area of Mobile IPv6 possible, we implemented an enhanced IPv6 testbed which is connected to the worldwide native "6net" infrastructure. As it can be seen in Figure 5, we built up a central core network where all subnetworks are attached to. In between each included network provider, we implemented WAN-Emulators that thwart all IPv6 packets transmitted. As our major aim was to create a very flexible network infrastructure, we put a single WAN-Emulator for each provider. So we are able to tune the link-delay individually, depending on the appropriate scenarios to be analyzed. Wireless LAN IEEE 802.11b and IEEE 802.11g are deployed in the overall infrastructure.

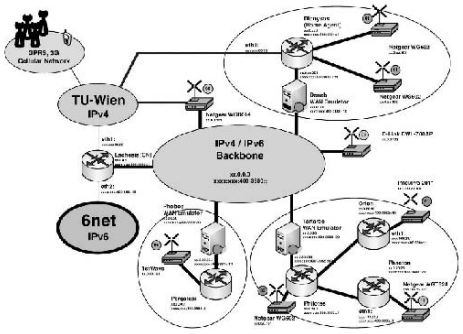


Figure 5: Mobile IP Testbed at TU-Vienna.

Three independent network operator domains were deployed, whereas one includes the Home Agent for the mobile node experiments. Furthermore, another network operator domain includes some kind of hierarchical structure in order to be able to do a performance comparison with the alternative HMIPv6 approach. All hosts including mobile nodes, correspondent nodes and the routers within each provider's area have RedHat Linux 8.0 installed with Kernel 2.4.22. For the MIPv6 basis functionality we utilized MIPL 1.0, provided by Helsinki University of Technology (HUT).

The Linux driver for all WLAN activities is based on the HostAP project, which seems to be the most flexible environment for making link-layer triggering realizable in a very fast manner. HostAP provides a general Linux driver for all PRISM2/2.5/3 based Wireless LAN cards. The results of an initial link-layer trigger optimization can be seen in Figure 3 and 4. These measurements are deployed by skipping the active-scanning mechanism within each handover.

5 PERFORMANCE EVALUATION

In this section we present initial results for a verification of the implemented IPv6 mechanisms and furthermore results based on our real-world Mobile IPv6 network infrastructure. For all measurements we derived average-values from about 1000 samples for each point in the graphics. This helps us to get significant and serious results for comparing of standard Mobile IPv6 to the enhanced FMIPv6 approach.

The first graph presents the difference in between communication with and without Route Optimization. The results of the end-to-end delay, depending on various link-delays, are presented in Figure 6.

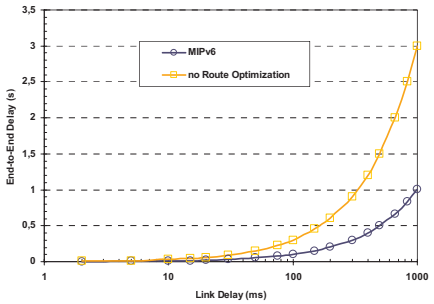


Figure 6: Route Optimization Impact on End-to-End Delay.

Figure 7 depicts the dependence of the handoff latency (foreign link – foreign link) on the variance of sending Router Advertisements. Obviously, the handoff latency falls off as Router Advertisement messages are sent more frequently.

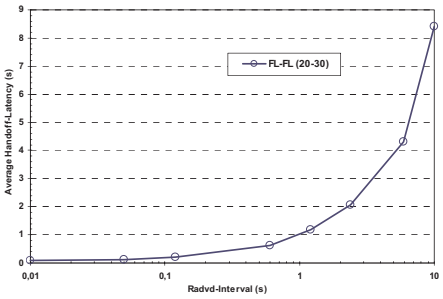


Figure 7: Handoff Latency for varying Router-Adv. Interval.

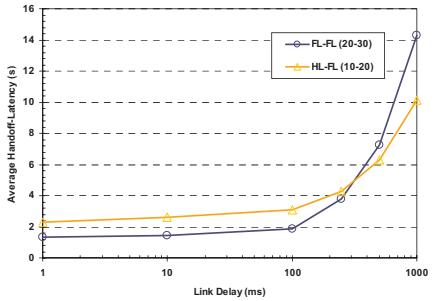


Figure 8: Average Handoff-Delay for Basic Mobile IPv6.

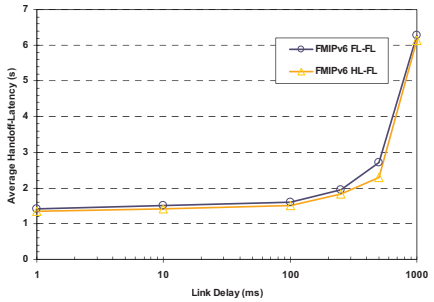


Figure 9: Average Handoff-Delay for Fast Mobile IPv6.

The results in Figure 8 and Figure 9 present the average handoff latency with dependence on the link-delay between different networks. Here we directly compared basic Mobile IP to the Fast MIPv6 approach.

As already assumed from the Fast Mobile IPv6 approach, the packet loss during a handover between different network providers is decreased to a minimum compared to basic Mobile IPv6. Figure 10 depicts the packet loss results for an Iperf- generated UDP-data stream of 160 kbit/s in between the mobile node and its correspondent node. As illustrated in Figure 5 the Correspondent Node is placed near the core network.

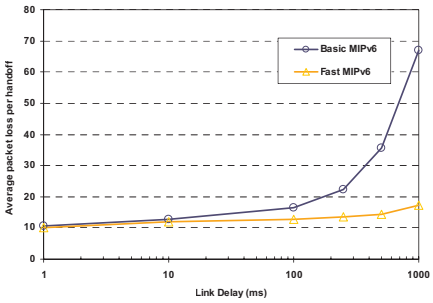


Figure 10: Average Number of Packet Loss during HO.

6 SIMULATIVE COMPARISON

For a deeper understanding as well as for a more general evaluation of Mobile IPv6 in an environment with many users, the use of simulations is indispensable. We performed a simulative comparison of baseline Mobile IPv6 and the Fast Handoff approach in an wireless LAN based scenario, comprising 4 independent operator domains with 10 home users per access router. Even if the focus is on the evaluation of MIPv6 bases protocols, we also include the impact of a shared-link environment based on IEEE 802.11b.

6.1 Simulation Scenario

For the performance study of MIPv6 we decided to evaluate a basic scenario which is simple enough to get results in a reasonable time but also complex enough to get an expressive feeling for real-world provider scenarios. The studied scenario (see Figure 11) is composed of a group of Correspondent Nodes, one for each Mobile Node, connected to one central router (CR) through the IPv6 backbone. Each access router (AR) represents a different IP subnet and acts as a home agent for 10 mobile nodes. All Mobile Nodes are located at their home link when the simulation starts. Either the distance in between the ARs and also the transmitted signal-power are chosen in a way to create overlapping coverage areas for enabling seamless movement in between the various domains (see Figure 12).

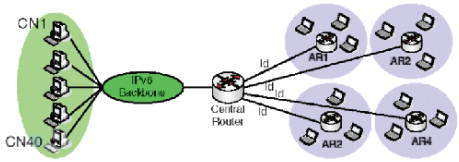


Figure 11: Mobile IPv6 – Simulation Scenario.

The random way-point mobility model is used for all Mobile Nodes, which is best suited for realistic user movement. Connectivity for each Mobile Node is provided by IEEE 802.11 using 2 Mbit/s and DCF and traffic is assumed to be UDP with 40 kbit/s constant bit rate.

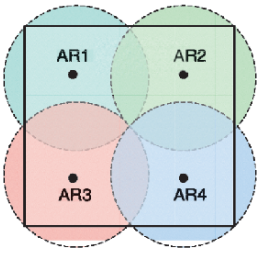


Figure 12: Access Router Range Topology.

For all our experiments we used the ns-2 (Ns2, 1998) simulation tool, whereas the MOBIWAN add-on by Thierry Ernst (Ernst, 2002) is deployed to get basic MIPv6 functionality into the simulator. Further essential MIPv6-specific software code was adopted from a MIPv6 simulation environment by NEC Europe in Germany.

6.2 FMIPv6 Simulation Results

In this section, we present the results of our ns-2 simulative comparison of baseline Mobile IPv6 and the enhanced Fast Handover mechanism.

Figure 13 presents the comparison of the handoff latency obtained during basic Mobile IPv6 handoff with the latency resulting from a Fast Mobile IPv6 handover. The simulation results show that similar to the performance measurements in Section 5 we also achieve some latency-related advantage for scenarios with a huge number of concurrent moving users.

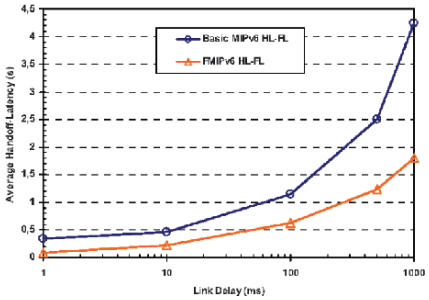


Figure 13: Average Handoff Latency Comparison.

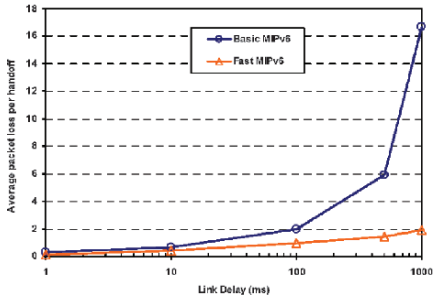


Figure 14: Average Packet-Loss per Handover.

With a reduced latency also the packet loss during the handoff can be reduced consequentially for the Fast Handoff approach. This behavior, similar to our testbed results from Section 5, is demonstrated in Figure 14.

7 CONCLUSIONS AND FUTURE WORK

With this work we presented a first evaluation and simulative results of a Fast Mobile IPv6 handover approach for wireless LAN based networks. Our evaluation showed that a client based fast handover approach can be suitable to improve WLAN handovers for real-time traffic and enables better mobility management support in IEEE 802.11 based wireless LANs. In the near future we will investigate on hierarchical approaches for IPv6 and other smart solutions with improved handoff latency performance and reduced signaling overhead.

ACKNOWLEDGEMENTS

Part of this work has been performed within the project "WISQY - Wireless InterSystem Quality-of-Service" at the Telecommunications Research Center Vienna (ftw) and has been funded in the framework of the Austrian Kplus Competence Center Programme. We would also like to thank Uschi Christalon-te Kock from NETGEAR for the great support within our research and Richard Menedetter for his assistance during the measurements.

REFERENCES

Proceedings Papers:

Johnson, D.B., Perkins, C.E., and Arkko, J., 2004. *Mobility Support for IPv6,* RFC 3775, IETF Network Working Group.

Koodli, R., 2004. *Fast Handovers for Mobile IPv6,* Internet-Draft, IETF Mobile IP Working Group, <draft-ietf-mipshop-fast-mipv6-02.txt>.

IEEE, 1999. *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,* IEEE Standard 802.11.

Velayos, H., and Karlsson, G., 2003. *Techniques to Reduce IEEE 802.11b MAC Layer Handover Time,* KTH Technical Report, ISSN 1651-7717, Stockholm, Sweden.

Velayos, H., and Karlsson, G., 2004. *Techniques to Reduce IEEE 802.11b Handoff Time,* IEEE ICC, Paris, France.

Mishra, A., Shin, M., and Arbaugh, W., 2002. *An Empirical Analysis of the IEEE 802.11 MAC Layer*, Technical Report CS-TR-4395, University of Maryland, Department of Computer Science.

Jordan, N., Fleck, R., and Ploninger, C., 2003. *Fast Handover Support in Wireless LAN based Networks*, Fifth IFIP-TC6 International Conference on Mobile and Wireless Communications Networks, Singapore, Singapore.

Ns2, 1998. *Network Simulator (ns),).* version 2, UC Berkeley, http://www.isi.edu/nsnam/ns.

Ernst, T., 2002. *Mobiwan*: ns-2 extensions to study mobility in Wide-Area IPv6 Networks, MOTOROLA Labs Paris, http://www.inrialpes.fr/planete/mobiwan.

Costa, X.P., Schmitz, R., Hartenstein, H., and Liebsch, M., 2002. *A MIPv6, FMIPv6 and HMIPv6 Handover Latency Study Analytical Approach*, Network Laboratories, NEC Europe Ltd.

Books:

Gast, M.S., 2002. *802.11 Wireless Networks, The Definitive Guide*. Sebastopol, CA: O'Reilly & Associates, Inc.

Intersil Corporation, 2001. *PRISM Driver Programme's Manual, Development Information*, Version 2.10, (For Distribution Under NDA Only).

CDMA2000 1X CAPACITY DECREASE BY POWER CONTROL ERROR IN HIGH SPEED TRAIN ENVIRONMENT

Simon Shin, Tae-Kyun Park, Byeung-Cheol Kim, and Yong-Ha Jeon
Network R&D Center, SK Telecom,
9-1, Sunae-dong, Bundang-gu, Sungnam City, Gyunggi-do, South Korea
Email: petites_flammes@nate.com

Dongwoo Kim
School of Electrical Engineering & Computer Science, Hanyang Univ.
1271 Sa-dong, Ansan, Kyungki-do 425-791, South Korea
Email: dkim@hanyang.ac.kr

Keywords: CDMA2000 1X, Doppler shift, capacity, power control, Korea Train Express.

Abstract: CDMA2000 1X capacity was analysed in the high speed train environment. We calculated the power control error by Doppler shift and simulated bit error rate (BER) at the base station. We made the interference model and calculated the BER from lower bound of power control error variance. The reverse link BER was increased by high velocity although there was no coverage reduction. Capacity decrease was negligible in the pedestrian (5 km/h), urban vehicular(40 km/h), highway and railroad(100 km/h) environment. However, capacity was severely reduced in high speed train condition(300 km/h and 350 km/h). Cell-planning considering capacity as well as coverage is essential for successful cellular service in high speed train.

1 INTRODUCTION

Cellular mobile telephone and data communication services are very popular. Cellular service is usable in anywhere, even though tunnel, sea, and underground places. Railroads and highways are main service area of cellular service because many people move through them. Many countries adopts the high speed train for transportation capability and convenience for example TGV in France, ICE in Deutschland. Korea also constructs the Korea Train Express (KTX) between Seoul and Pusan. KTX travels by 300 km/h speed and will be upgraded by 350 km/h speed.

High speed mobility can damage pilot acquisition, steady-state demodulation, code tracking, and power control. We experimented the CDMA2000 1X service quality using channel simulator in Test-bed network. There was no quality degradation such as MOS(mean of score), data throughput, call drop, and call fail. We got the same results in the KTX. Call origination, response, data download, and data upload was successful in the train with 300 km/h velocity. Received power, transmitted power, and pilot chip energy to interference ratio (Ec/Io) of mobile station were not correlated with the mobile velocity. We could serve successfully the CDMA2000 1X in the KTX by existing cellular network.

Experimental results confirm that coverage is not reduced by speed mobility. However, a few base station and frequency assignment (FA) is established in rural area. It makes the capacity shortage. One KTX train is composed of 20 cars and one car seats 64 persons. It means more than 2000 calls can connect one base station when two opposite trains are in coverage of same base station. In the dense urban area, heavy traffic cell is split for dividing traffic. This method cannot be used for KTX because traffic cannot be split in one train. We must increase the FA for capacity improvement. Adjacent cell uses same FA for smooth handoff. If one cell increases the FA number, adjacent cells also increase the FA number. Therefore we have to estimate the capacity of one FA for efficient network investment.

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 227–231.

Velocity does not influence the coverage, but it does not mean that capacity is not varied by velocity. Network performance can be changed with user number because CDMA is the system limited by interference. CDMA performance is influenced by multi-user interference. CDMA capacity is limited by Walsh code on the forward link and by cell loading on the reverse link. Cell loading is the received power increase by mobile station at the base station receiver. High cell loading increases the noise level at the base station receiver and it degrades the bit energy to noise ratio (Eb/Nt). Low Eb/Nt increases the transmitted power of mobile station by power control and high transmitted power degrades Eb/Nt repeatedly. CDMA2000 1X uses the fast power control for corresponding to variable radio channel. High velocity of mobile causes the power control error due to Doppler shift and fast radio channel environment variation. This paper analysed the capacity decrease by power control error. Section 2 will derive the calculation of power control error and Section 3 will show the simulation results.

2 POWER CONTROL ERROR SIMULATION

2.1 Interference Model

We starts making interference model by followed assumption.

- User *o* who is analysis target has the appropriate received signal strength 0 dB.
- There are K+*1* users including target user *o.*
- *K+1* users are served in same cell.
- *K+1* users are riding the KTX.
- There is no other user who is not riding KTX in the analysed cell. That is to say, all of users in the analysed cell are riding KTX.
- Other cell user *K* interference is assumed in-cell user *gK*.
- There is no error in the power control procedure of the other cell user.

Received signal in base station from interference user *i* is as follows:

$$S_i = m_i^{+}\varepsilon_i \quad \text{(dB)} \tag{2.1}$$

m_i is the signal strength without power control error. ε_i is the random variable due to power control error and has log-normal distribution with 0 dB expected value and σ_i^2 dB variance. Expected value and variance of random variable S_i measured by Watt-unit are as follows.

$$E[S_i] = e^{\beta m_i + \frac{1}{2}\beta^2\sigma_i^2} \tag{2.2}$$

$$Var(S_i) = e^{2\beta m_i + \beta^2\sigma_i^2}\left(e^{\beta^2\sigma_i^2} - 1\right) \tag{2.3}$$

$\beta$is (ln10)/10 in equation 2.2 and 2.3. Received signal strength of user *o* is as equation 2.4.

$$S_o = \varepsilon_o \quad \text{(dB)} \tag{2.4}$$

In equation 2.4, it is assumed ε_o has a log-normal distribution with variance σ_o^2 (dB).

Received power from other cell users must be considered to analyse multi-cell environment. Interference from other cell is represented by in-cell interference. If there are *K* users in the each cell and average propagation loss ratio of in-cell and other cell is *g*, other cell interference is *gK*. Standard deviation of received signal is the control function of base station transmitted power because of fading. Total interference considering both in-cell and other cell interference is as equation 2.5

$$I = \sum_{i=1}^{K} S_{ei} + \sum_{i=1}^{gK} S_{oi} \tag{2.5}$$

S_{ei} is signal power of in-cell user *i* with standard deviation σ_{ei}. S_{oi} is signal power of other cell user *i* with standard deviation σ_{oi}. Therefore, expected value and variance of *I* is as follows

$$E[I] = \sum_{i=1}^{K} E[S_{ei}] + \sum_{i=1}^{gK} E[S_{oi}] \tag{2.6}$$

$$Var(I) = \sum_{i=1}^{K} Var(S_{ei}) + \sum_{i=1}^{gK} Var(S_{oi}) \tag{2.7}$$

Probability density function (PDF) of *I* can be approximated by log-normal random variable ξ, because *I* is sum of independent log-normal random variables. Expected value and variance of ξ is as follows

$$E[\xi] = e^{m_K + \frac{1}{2}\sigma_K^2} \tag{2.8}$$

$$Var(\xi) = e^{2m_K + \sigma_K^2}\left(e^{\sigma_K^2} - 1\right) \tag{2.9}$$

It is assumed power control error variance of in-cell users and other cell is σ_e^2 and σ_o^2, respectively. Expected value and variance of *I* is defined with expected value and variance of ξ using Wilkinson's Method.

$$Ke^{\beta m+\frac{1}{2}\beta^2\sigma_\varepsilon^2}+gKe^{\beta m+\frac{1}{2}\beta^2\sigma_o^2}=e^{m_K+\frac{1}{2}\sigma_K^2} \quad (2.10)$$

$$Ke^{2\beta m+\beta^2\sigma_\varepsilon^2}\left(e^{\beta^2\sigma_\varepsilon^2}-1\right)+gKe^{2\beta m+\beta^2\sigma_o^2}\left(e^{\beta^2\sigma_o^2}-1\right)=e^{2m_K+\sigma_K^2}\left(e^{\sigma_K^2}-1\right) \quad (2.11)$$

We get m_k and σ_k from equation 2.10 and 2.11.

$$\sigma_K^2=\ln\left[\frac{e^{\beta^2\sigma_\varepsilon^2}\left(e^{\beta^2\sigma_\varepsilon^2}-1\right)+ge^{\beta^2\sigma_o^2}\left(e^{\beta^2\sigma_o^2}-1\right)}{K\left(e^{\frac{1}{2}\beta^2\sigma_\varepsilon^2}+ge^{\frac{1}{2}\beta^2\sigma_o^2}\right)^2}+1\right] \quad (2.12)$$

$$m_K=\ln K+\ln\left(e^{\frac{1}{2}\beta^2\sigma_\varepsilon^2}+ge^{\frac{1}{2}\beta^2\sigma_o^2}\right)-\frac{1}{2}\sigma_K^2 \quad (2.13)$$

2.2 Mean BER Calculation

Bit energy to interference ratio of BPSK having bandwidth W is given by equation 2.14.

$$\frac{E_b}{N_0+I_0}=\frac{S_{eo}/R_b}{N_0+\frac{1}{W}\left(\sum_{i=1}^{K}S_{ei}+\sum_{i=1}^{gK}S_{oi}\right)} \quad (2.14)$$

N_o is the power spectral density of background noise and R_b is the bit rate. BER of BPSK is given by

$$P_e=Q\left(\sqrt{2\frac{E_b}{N_0+I_0}}\right) \quad (2.15)$$

The received signal is log-normal random variable due to imperfect power control and BER is given by

$$P_e=Q\left(e^{\gamma}\right) \quad (2.16)$$

γ is the Gaussian random variable and its mean and variance is given by

$$m_\gamma=\frac{1}{2}\left(\ln 2\frac{W}{R_b}-m_K\right) \quad (2.17)$$

$$\sigma_\gamma^2=\frac{1}{4}\left(\beta^2\sigma_e^2+\sigma_K^2\right) \quad (2.18)$$

With mean γ, mean BER is calculated by

$$\overline{P}_e=\int_0^\infty Q\left(e^{\gamma}\right)g(\gamma)d\gamma \quad (2.19)$$

Equation 2.19 is approximated to equation (2.20) using the expansion of central difference.

$$\overline{P}_e\approx\frac{2}{3}Q\left(e^{m_\gamma}\right)+\frac{1}{6}Q\left(e^{m_\gamma+\sqrt{3}\sigma_\gamma}\right)+\frac{1}{6}Q\left(e^{m_\gamma-\sqrt{3}\sigma_\gamma}\right) \quad (2.20)$$

2.3 Variance of Power Control Error

Each base station controls the received power from mobile station in cellular CDMA network. There is four main error factor when tracking the received power. They are the quantum(σ_q), decoding(σ_d), measurement(σ_m), and propagation delay(σ_p) error. Generally, σ_m and σ_p is much larger than σ_q and σ_d. If error factor is statistically independent,

$$\sigma_e^2=\sigma_m^2+\sigma_p^2+\sigma_q^2+\sigma_d^2 \quad (2.21)$$

$v$is the maximum velocity of mobile. f_c is the carrier frequency. c is the light velocity. We assume the bandwidth is narrow enough to neglect bandwidth. Maximum Doppler shift is

$$f_d=f_c v/c \quad (2.22)$$

Propagation delay is $2d/c$ for closed loop power control. d is the distance between mobile and base station. Therefore, processing delay is

$$T_p<\frac{\alpha}{f_d}-\frac{2d}{c} \quad (2.23)$$

We define the p_m is measured received power in dB scale. p_0 is the received power in linear scale and p_m is $10\log p_0$. If σ_{m1}^2 is variance of natural logarithm($\ln p_0$) of p_0

$$\sigma_m^2=(10\log e)^2\sigma_{m1}^2 \quad (2.24)$$

It is assumed that fluctuation of received power can be neglected during measurement time T_m. Measurement period T_m is the main factor of processing delay T_p. Power control error is due to multi-user interference and white Gaussian noise. We assume the power control makes received power from any mobile to be almost same. It enables us to determine the lower bound. Multi-user interference is modelled Gaussian process that increases one-sided power spectral density N_o to N_t.

$$N_t=N_0+\frac{p_0}{W}K(1+g) \quad (2.25)$$

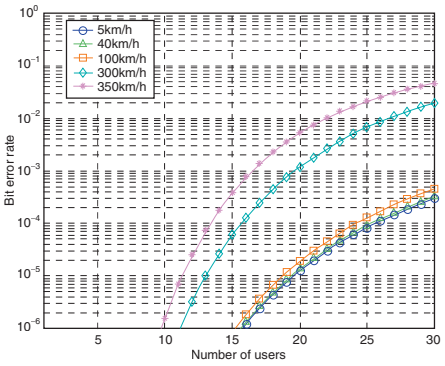


Figure 1(a): $g = 0.3$ and $\sigma_p^2 + \sigma_q^2 + \sigma_d^2 = 0.5(\text{dB})^2$.

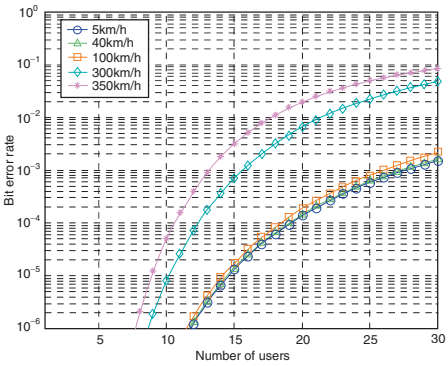


Figure 2(a): $g = 0.634$ and $\sigma_p^2 + \sigma_q^2 + \sigma_d^2 = 0.5(\text{dB})^2$.

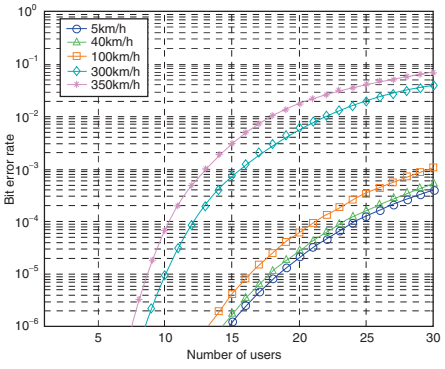


Figure 1(b): $g = 0.3$ and $\sigma_p^2 + \sigma_q^2 + \sigma_d^2 = 0.5(\text{dB})^2$.

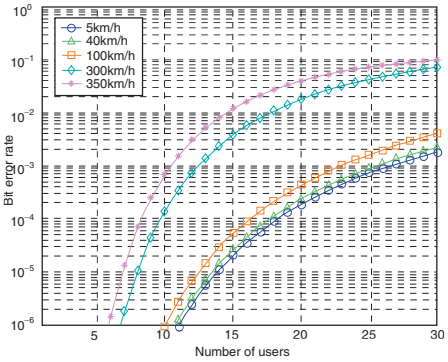


Figure 2(b): $g = 0.634$ and $\sigma_p^2 + \sigma_q^2 + \sigma_d^2 = 0.5(\text{dB})^2$.

p_o is *K-1* interference signal power. *W* is the receiver bandwidth. g is the interference constant that represents other cell interference. Received signal is $\sqrt{p_0}s(t)$.

$$\sqrt{p_0}s(t) = \exp\left(\frac{y}{2}\right)s(t) \tag{2.26}$$

y is $\ln p_o$ in equation 2.26. Cramer-Rao bound provides lower bound of $\ln p_o$ variance. Equation 2.27 is obtained from this lower bound and equation 2.26.

$$\sigma_{ml}^2 \geq \left\{\frac{2}{N_t}\int_0^{T_m}\left[\frac{\partial}{\partial y}e^{y/2}s(t)\right]^2 dt\right\}^{-1} \tag{2.27}$$

using equation 2.24 and 2.25

$$\sigma_m^2 \geq \frac{200(\log e)^2}{T_m}\left[\frac{N_0}{p_0} + \frac{K(1+g)}{W}\right] \tag{2.28}$$

If we define $T_1 = T_p - T_m$ and use equation 2.23, we obtain equation 2.29 from 2.21

$$\sigma_{low}^2 \triangleq 200(\log e)^2\left(\frac{\alpha}{f_d}-\frac{2d}{c}-T_1\right)^{-1}\left[\frac{N_0}{p_0}+\frac{K(1+g)}{W}\right]+\sigma_p^2+\sigma_q^2+\sigma_d^2 \quad (2.29)$$

3 SIMULATION RESULTS

We simulated the BER of received signal at base station when user number was 5 ~ 30 persons. Simulation circumstance was assumed to be pedestrian(5 km/h), urban vehicle(40 km/h), highway and railroad(100 km/h), KTX(300 km/h), and upgraded KTX(350 km/h).

BER was calculated from equation 2.20. m_r and σ_r of equation 2.20 was calculated from equation 2.17 and 2.18. m_k and σ_k was calculated from equation 2.12 and 2.13. We used the lower bound of equation 2.28 when calculating equation 2.12 and 2.13. Figure 1 and 2 shows user number and BER when mobile speed is 5 km/h, 40 km/h, 100 km/h, 300 km/h, and 350 km/h. We assumed R_b = 4.8 kbps, W = 1.2288 MHz, d = 4 km, α= 0.1, T_1 = 100 us, N_o/p_o = 5 us, σ_o^2 = 3.9 (dB)^2. Figure 1 and 2 shows increase of mobile speed degrades the BER of received signal. This means the reverse link capacity decreases to maintain the quality of service. Figure 1 and 2 shows the result when interference constant g is 0.3 and 0.634, respectively. Larger interference constant increases the receiver sensitivity with user number. (a) and (b) of each Figure shows the results in the case of $\sigma_p^2+\sigma_q^2+\sigma_d^2 = 0.5(\text{dB})^2$ and $\sigma_p^2+\sigma_q^2+\sigma_d^2 = 1.5(\text{dB})^2$. High speed enlarges the Doppler shift in equation 2.22. Doppler shift increases the lower bound of error variance in equation 2.29. This increases the BER and degrades the quality of service.

In urban vehicular (40 km/h) condition, BER increase by Doppler shift was negligible. BER degradation was not severe even though highway and railroad (100 km/h) condition. We could plan the cellular network assuming constant capacity with mobile speed before KTX service. However, BER was dramatically increased in KTX circumstance. User number in KTX was limited to 17 ~ 26 persons to maintain BER lower than 1%.

4 CONCLUSION

We measured the coverage of CDMA2000 1X network experimentally and simulated the capacity in KTX condition. Although coverage was not decreased, capacity was reduced severely in high mobile speed of 300 km/h. We don't have to consider the mobile velocity in cell-planning because capacity reduction is negligible in highway and railroad. However, capacity is severely reduced in KTX for its high velocity. We must consider the number of passenger carried by KTX when opposite train is met. Cell-planning without considering capacity can make the burst error in high traffic intensity. It causes not only quality degrade but also call drop. We must consider capacity as well as coverage for cellular network planning.

REFERENCES

Viterbi, A.M., Viterbi, A.J., and Zehavi, E., "*Other-cell interference in cellular power-controlled CDMA*," IEEE Trans. on Communications, vol. 42, No. 2/3/4, pp. 1501-1504, Feb./Mar./Apr. 1994.

Hashem, B., and Sousa, E., "*Increasing the DS-CDMA system reverse link capacity by equalizing the performance of different velocity users*," Pro. IEEE ICC'98, Atlanta, Georgia, USA.

Beaulieu, N.C., Abu-Dayya, A.A., and McLane, P.J., "*Estimation the distribution of a sum of independent lognormal random variables*," IEEE Trans. on Communications, vol. 43, No. 12, pp. 2869-2873, Dec. 1995.

Pursley, M.B., "*Performance evaluation for phase-coded spread spectrum multiple-access communication – part I : system analysis*," IEEE Trans. on Communications, vol. COM-25, No. 8, pp. 795-799, Aug.1977.

Holtzman, J.M., "*A simple, accurate method to calculate spread-spectrum multiple-access error probabilities*," IEEE Trans. on Communications, vol. 40, No. 3, pp. 461-464, Mar. 1992.

Lee, C.C., and Steele, R., "*Closed-loop power control in CDMA systems*," IEE Proc.-Commun., vol. 143, pp. 231-239, Aug. 1996.

McDonough, R.N., and Whalen, A.D., *Detection of Signals in Noise*, 2nd ed. San Diego : Academic Press, 1995.

Torrieri, D., "*Uplink capacity of a CDMA network*," internal communication.

UGSP: AUTHENTICATION BASED SECURE PROTOCOL FOR AD-HOC NETWORKS

Neelima Arora
Tata Institute of Fundamental Research
Mumbai, India
Email: neel@iitbombay.org

R. K. Shyamasundar
Tata Institute of Fundamental Research
Mumbai, India
Email: shyam@tcs.tifr.res.in

Keywords: Ad-hoc networks, key distribution, security

Abstract: A wireless ad-hoc network is a collection of mobile nodes with no fixed infrastructure. Security in such networks poses serious challenges due to (i) the network connectivity could be intermittent and hence on-line authentication is not guaranteed, and (ii) susceptible to wide range of attacks due to broadcast communication and large scale number of users. In this paper, we propose a security protocol, called UGSP, for wireless ad-hoc networks using a tamper-proof hardware. We show that the proposed protocol fits well with the resurrecting duckling security paradigm (Stajano and Anderson, 1999). Once the hardware is imprinted for authentication, UGSP is robust to man-in-the-middle attack, passive eavesdropping, active impersonation attacks ensuring source authentication, data confidentiality and data integrity for communication amongst nodes with identically configured hardware. The system is amenable to dynamic addition of new members whose hardware has also been imprinted with authentication information. We provide a comparative evaluation of UGSP with other approaches and show that UGSP is scalable and cost-effective.

1 INTRODUCTION

Ad-hoc networks do not have fixed infrastructure such as base station or mobile switching centers. Mobile nodes, which are within the range of each other, communicate directly while those that are far apart rely on other nodes to forward messages as routers. Node mobility in the network causes frequent changes of the network topology.

The salient features of ad-hoc networks pose both challenges and opportunities in achieving security goals characterized by attributes like availability, confidentiality, integrity, source authentication and non-repudiation (Zhou and Haas, 1999). Nodes, roaming in hostile environment (e.g., a battlefield) with relatively poor physical protection, have non-negligible probability of being compromised. Therefore, we should not only consider malicious attacks from outside a network, but also take into account the attacks launched from within the network by compromised nodes. We envision ad-hoc networks to be formed by nodes without any prior contact, trust or authority relation. This precludes any pre-distributed symmetric keys or a reliable (external) PKI supported across all nodes. This issue has been largely ignored until now; most protocols assume that the key-distribution has already taken place. Further, security mechanisms should be scalable to handle large networks.

1.1 Current State of Ad-Hoc Network Security

One of the largely investigated areas of ad-hoc network security research is devoted to secure routing protocols (Toh, 2001)(Royer and Toh, 1999), that form an essential component of security in ad-hoc networks (Khaili and Arbaugh, 2002). However, most of the routing schemes known neglect the crucial challenge in ad-hoc security: key establishment and key distribution. Protocols such as ARAN, Ariadne(Hu and Perrig, 2002), SPINS(Perrig et al., 2002b), TESLA(Perrig et al., 2002a), SEAD (Hu et al., 2002) and SRP (Papadimitratos and Haas, 2002) all assume the pre-existence and pre-sharing of secret and/or public keys for all the nodes. In other words, key management and key distribution in an ad-hoc networks has been left a wide open problem. Recently some approaches for key distribution in ad-hoc networks have been proposed (Zhou and Haas, 1999; Bobba et al., 2002; Khalili and Arbaugh, 2003). Note

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 233–238.

that a mechanism is needed wherein we can accommodate the new trust scenarios in ad-hoc networks.

1.2 Our Approach

In this paper we assume that both active and passive attackers are present in the environment. The attacker is allowed to watch regular runs of the protocol between the two communicating nodes and also send arbitrary messages to the parties. We also assume that the attacker has sufficiently large computational power.

In this paper, we describe a protocol called UGSP (User Group Security Protocol) for communication amongst a dynamic user group (DUG), that is resilient against the above mentioned attacks while maintaining all the necessary attributes (confidentiality, integrity and authenticity). Even if the node gets compromised, it should not allow the attacker to gain access to any useful secret of the network. We see examples of DUG in our day to day life such as employees of a company or all the mobile cell-phone users of a particular network.

Rest of the paper is organized as follows: section 2 gives the system architecture required for the protocol, section 3 details the protocol, analysis of the protocol is done in section 4, section 5 gives the implementation details and section 6 contains the conclusion, future work and generalizations.

2 SYSTEM ARCHITECTURE FOR THE UGSP

To make UGSP as generic as possible, we have designed the protocol minimizing the energy cost. Most previous work on secure ad-hoc network relies on asymmetric cryptography for establishing the security parameters every time. However, computing such signatures on resource-constrained nodes is expensive and hence may not be the ideal solution. A protocol with shared secret is the most generic option, as it is not expensive both in terms of bandwidth and computation.

The system has the following three components:

(1) Communication Link: Our system does not assume any special characteristic of the communication link and will work over any form of communication system such as ethernet, bluetooth, 802.11, optical fiber etc.

(2) Nodes: Every node that participates in the ad-hoc network is assumed to have the structure shown in Figure 1. That is, it has

- Processing and storing capabilities as demanded by the protocol,

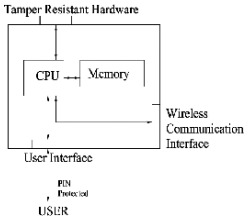


Figure 1: A valid node in the network.

- Tamper Resistant Hardware: the capabilities of this hardware are described in the next paragraph,
- Interface for the tamper resistant hardware.

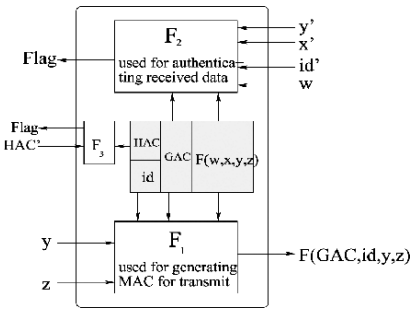


Figure 2: Hardware on the node.

The structure of the ideal tamper proof hardware (TPH) for UGSP is shown in Figure 2. The salient features of the TPH are given below:

- a universally unique, unalterable id
- Write only memory for group access code (GAC) and hardware access code (HAC), attempt to overwrite HAC destroys GAC as well.
- non invertible pseudo random hash function $F(w, x, y, z)$ embedded in the hardware (Lamport, 1981)
- provides service of three functions F_1 (used to generate message digest for transmission), F_2 (used for authenticating the message digest received from the sender) and F_3 (used for local authentication of the HAC) defined below:
  $F_1(y, z) = F(GAC, id, y, z)$
  $F_2(w, id', y', x')$
  $= TRUE \ \ if \ \ w = F(GAC, id', x', y');$
  $= FALSE \ \ otherwise$
  $F_3(HAC', HAC) = TRUE \, if \, HAC = HAC'$
  $= FALSE \ \ otherwise$
- Tamper resistance of the hardware is interpreted as:

  - Codes of functions $(F(w,x,y,z), F_1, F_2, F_3)$ and id cannot be accessed by any one, and
  - For F_1, id and GAC are implicit (or hidden) parameter, for F_2, GAC is implicit parameter and HAC is implicit parameter for F_3.
- Password Enabled Access: the use of the hardware will be restricted by a HAC. Only after a user has entered the HAC correctly, he will be able to use the TPH. This ensures source authenticity at the user level.

Imprintable TPH can be obtained in the market and the user can download the image for the specific application over a secure channel to imprint the TPH. (3) User Group Administrator: We envisage the presence of a User Group (UG) Administrator with the following roles:

- Assign a GAC for the UG
- Develop and maintain image for imprinting TPH
- Provide image and configure the TPH on request by a user on getting necessary information.

The role of UG administrator is limited to configuring the nodes and does not have any role during the interaction between the end nodes.

3 UGSP: USER GROUP SECURITY PROTOCOL

The TPH token is not integrated with the mobile host, but the user carries the token with himself.
The protocol consists of the following steps:

1. First, every node that is allowed into the network is initialized by the group administrator by writing the HAC (user specific) first and then the GAC (which is same for all nodes) in the write only memory of the tamper resistant hardware. HAC is written before GAC because write to HAC overwrites GAC also. The write-only memory is interpreted as follows: once the secret is written into the memory, no one can read the secret. That is, a node can be a part of a DUG ad-hoc network provided it has the TPH with the same GAC.
2. The user will carry the configured TPH with him as a token. In order to use the hardware, the user will be required to interface it with his mobile device and to enter the HAC correctly, which will be locally authenticated within the hardware. Only after successful authentication, he can use the hardware for the subsequent steps.
3. Communicating nodes generate 1024 bit RSA key pair and exchange their public keys using the MAC generated by the TPH.
4. The sender now generates a 64 bit DES symmetric key, encrypt it using the receiver Public Key, append a MAC generated by the hardware and transmit. The receiver also confirms that he has received the correct symmetric key by sending the symmetric key encrypted with the sender Public Key.

Note that due to the underlying tamper proof hardware, the following properties are satisfied:

- No user has any control over implicit inputs to the functions F_1, F_2 and F_3.
- When a node wants to transmit data, it gives the data to the TPH which in turn evaluates the MAC using function F_1 and transmits. Note that the user cannot change the first two inputs of function F_1, i.e. the GAC and id of the node (they are taken automatically from the memory of the hardware). Because the id cannot be changed, active impersonation is not possible in the network.
- Now consider a node receiving data and the corresponding MAC. The node now needs to authenticate the MAC against the data, id of the sender node and the GAC. For this purpose, the node passes the received data and the sender id to the TPH, which computes the function F_2 taking the GAC stored in its memory. If the result of F_2 is $TRUE$ then the node will accept the data otherwise the node realizes that there is some mismatch in data, id or GAC, and rejects the packet and terminates the communication.

3.1 Formal Description

Notation

- Nodes *i* and *j* want to communicate with each other, where i is the sender and j is the receiver.
- *F(w, x, y, z)* is the pseudo random hash function implemented in the TPH, which takes in four input parameters w, x, y, z
- *id(k)* is the identity of the node *k*
- K_j is the public key of the node *j*
- $E_{K_j}(N)$ means *N* encrypted with K_j
- *N* denotes the $nonce$ which is unpredictable.
- $(k \rightarrow l : M)$ is interpreted as follows: node *k* is sending message, *M*, meant for node *l*
- K_{sy} is the symmetric session key established between the pair of communicating nodes
- (M_1, M_2) means message M_1 concatenated with M_2
- For ease of reference in the sequel, we shall refer the value of $F(w,x,y,z)$ as MAC.

In UGSP, even if the attacker is able to predict the $nonce$, he will not be able to learn any secret in the network unlike most other protocols (will become clear in the sequel).

Every time a node wants to transmit, it computes MAC using F_1 and transmit, and a receiving node to authenticate the received data will compute F_2 as mentioned above. With this background, we are ready to describe the steps in the protocol formally.

The protocol consists of two Phases. Phase I of the protocol consists of bootstrapping of the valid nodes of the network, while the second phase will be session specific and will happen as and when nodes want to communicate with each other. In more hostile environments, such as battlefield, where the chances of bootstrapped nodes getting compromised is higher, one can make bootstrapping also session specific.

Operational steps in the execution of UGSP:

$Phase\ I$: the network is being formed. In Step 1, the user gets authenticated by the TPH by the HAC. User inputs the HAC when prompted to do so, which will be compared with the HAC stored on the TPH. This can be thought of as if the smart card PIN is stored on the smart card and which can only be read by the smart card reading machine. Hence, user authentication can be done by a standalone ad-hoc smart card machine, not connected to the bank's back-end server, by comparing the PIN stored on the card with the PIN entered by the user. In Step 2, the communicating nodes, which have been authenticated in step 1, will exchange RSA keys.

$Phase\ II$: will be invoked when any node want to communicate with some other node in the network. During step 3, the two communicating nodes establish a symmetric key. We could directly use the private-public key pair established in step 3 for secure communication, but that will be bandwidth and computation expensive. Hence, we need to establish a symmetric key, which is accomplished after this step. We encrypt the data using the symmetric key and append a MAC computed using the data and the symmetric key and transmit in step 4. As highlighted already, UGSP uses the Public Key operations in a limited way as compared to other protocols.

PHASE I : **Bootstrapping**
Step 1 : **local authentication of HAC**
1.a : Node i : user i enters his HAC
1.b : Node j : user j enters his HAC

Step 2 : **exchange public keys**
2.a : $\mathrm{i} \rightarrow \mathrm{j} : \mathrm{F}(\mathrm{GAC}, \mathrm{id(i)}, \mathrm{N_1}, \mathrm{K_i}), \mathrm{N_1}, \mathrm{K_i}, \mathrm{id(i)}$
2.b : $\mathrm{j} \rightarrow \mathrm{i} : \mathrm{F}(\mathrm{GAC}, \mathrm{id(j)}, \mathrm{E_{K_i}(N_1)}, \mathrm{K_j}), \mathrm{N_2}, \mathrm{K_j}, \mathrm{id(j)}$
2.c : $\mathrm{i} \rightarrow \mathrm{j} : \mathrm{F}(\mathrm{GAC}, \mathrm{id(i)}, \mathrm{E_{K_j}(N_2)}, *)$

PHASE II : **Session specific**
Step 3 : **establish a shared symmetric key using the established public key**
3.a : $\mathrm{i} \rightarrow \mathrm{j} : \mathrm{F}(\mathrm{GAC}, \mathrm{id(i)}, \mathrm{N_3}, \mathrm{E_{K_j}(K_{sy})}), \mathrm{N_3}, \mathrm{E_{K_j}(K_{sy})}$
3.b : $\mathrm{j} \rightarrow \mathrm{i} : \mathrm{F}(\mathrm{GAC}, \mathrm{id(j)}, \mathrm{E_{K_{sy}}(N_3)}, \mathrm{E_{K_i}(K_{sy})})$
3.c : $\mathrm{i} \rightarrow \mathrm{j} : \mathrm{F}(\mathrm{GAC}, \mathrm{id(i)}, \mathrm{E_{K_{sy}}(id(i))}, *)$

Step 4 : **data transfer**
4.a : $\mathrm{i} \rightarrow \mathrm{j} : \mathrm{E_{K_{sy}}(Data)}, \mathrm{F}(\mathrm{id(i)}, \mathrm{data}, \mathrm{K_{sy}})$

Figure 3: Protocol for communication between Node i and Node j.

4 ANALYSIS OF THE PROTOCOL

In this section, we will analyze UGSP with respect to security assurance and scalability.

4.1 Security Analysis

In this section we will discuss possible attacks in ad-hoc network and analize the security assurance provided by UGSP as well as the approaches of TESLA and PKI based security system. We shall show robustness of our protocol with respect to outside attack and the risks of compromised nodes.

4.1.1 Attacks from Outside the Network

When we say a node is outside the network, we mean a node that does not have the TPH token configured with the GAC for that group, say GAC_1. In this case, either the attacker will fail to authenticate himself to the hardware token in the $Step\ 1$ itself (cf. Figure 3) or if the user knows the valid HAC for the TPH (which means that he is a valid user for some other similar network with group access code GAC_2) then he will generate the MAC corresponding to GAC_2. Such a MAC would fail to match with the MAC generated at the receiving end because of the different GAC. While if such a node wants to become a receiver, then it will fail to authenticate itself to the sender as it will not be able to form a valid $Packet\ 2.b$ (cf. Figure 3) because of a different GAC. Thus, such a node would fail to be a sender as well as a valid receiver in the network.

4.1.2 Attack from a Compromised Node

Here, the attacker is assumed to have the valid compromised node, i.e., a node with the TPH token configured with the group access code for that group. Here, again two level of attacks are possible:

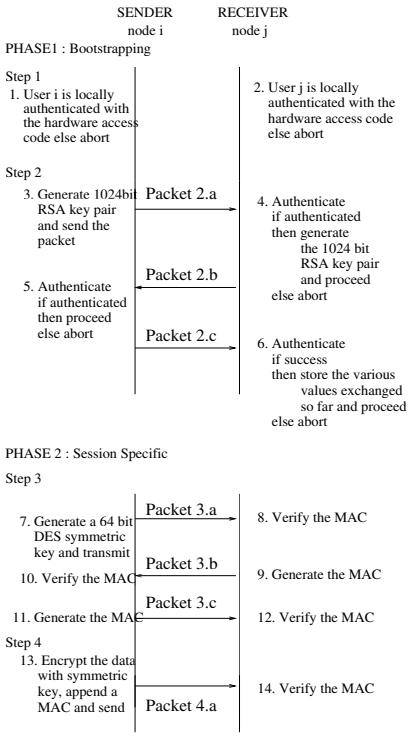


Figure 4: Algorithm at the nodes (Packets are labeled in Figure 3).

- *Attacker does not know the HAC:* The authentication will fail during Step 1 (cf. Figure 4). While if the attacker tries to overwrite the HAC, the GAC also gets overwritten (as mentioned above, TPH token is has been designed with such a feature) and then the case becomes similar to the one mentioned in Section 4.1.1.
- *Attacker knows the HAC:* In this case the attacker can send valid packets to the network, but with his own id. Active impersonation is not possible in the network. The attacker cannot simulate the hardware behavior and try to pass any other id to the model, because the group access code is not known outside the hardware. In such a scenario, if other nodes in the network come to know that a node is compromised, then they can block that particular node from any communication because when that node wants to participate in the communication, it can do so with its own identity, which is blocked.

In case of PKI or TESLA, when the node gets compromised the attacker would get the keys and will be able to actively impersonate the user. Even though the keys may be password protected but note that they would be accessible through brute-force methods (such as bit by bit reading of the hard disk), to which TPH is resistant.

4.2 Scalability

After the configuration of HAC and GAC at the hardware, the key establishment and data transfer etc can be initiated as per the ad-hoc network requirements. The nodes can form an ad-hoc network with other nodes in DUG securely without having to maintain any database, without going to trusted third party, without requiring any other authenticated channel for each data transfer. Thus, UGSP is suitable for ad-hoc wireless network security. Note that, the cost of TPH required for the purpose is available and is quite cheap. In other words, UGSP is quite scalable.

Let us consider the needs of TESLA and PKI-based security from the perspective of scalability. In TESLA, every time two nodes want to communicate, they will be required to exchange some information (like: T_{int}, key disclosure delay, key commitment to the key chain) over an authenticated channel. This is not possible in ad-hoc environment as it requires constant presence of a trusted third party.

Now, let us consider PKI-based security: Assume that each node eligible to join the network has been given a public-private key pair by PKI. Thus, when any node wants to communicate, it will send the certificate along with the data. For the other node to verify the certificate either it has to go to the trusted third party (which is not possible in ad-hoc network) or maintain its own database of public keys of the other nodes (which is not scalable). Due to the property that on-line authentication is needed, a typical PKI-based security is not feasible for ad-hoc network security.

5 IMPLEMENTATION DETAILS

We have implemented UGSP using $iButton$[1]. The $iButton$ is a computer chip enclosed in a 16mm stainless steel can available off-the shelf for less than ten dollars in retail. Note that the $iButton$ provides more functionality than what is needed for UGSP. We have used it as it is available off-the shelf. The steel button can be mounted virtually anywhere because it is rugged enough to withstand harsh environments, indoors or outdoors. Each $iButton$ has a unique and unalterable address that is laser etched onto its chip inside the can. In response to tampering, the $iButton$ would rather erase the key than reveal its secrets.

[1] www.ibutton.com

Each has an onboard 512-bit SHA-1 engine that can compute 160-bit MACs in less than 0.0005 seconds as compared to 0.5 seconds for a typical microcontroller. $iButton$ can be interfaced with a host system via serial/parallel port or USB. Our protocol has been tested with $iButton$. Mobility of the nodes was also emulated and multihop routing scenarios were evaluated against performance and energy cost. However, $iButton$ does not realize the function F_2 but indicates that the TPH described in the protocol can be developed at a much cheaper cost. We have completed the design of the required TPH and are in the process of testing it on an FPGA board.

6 DISCUSSIONS

To sum up, it can be seen that UGSP is resilient to all attacks on an ad-hoc network forming a DUG mentioned earlier. UGSP is based on mutual authentication rather than only the client authenticating to the server, or only the sender authenticating to the receiver. Our protocol provides dual security since we are using a TPH token and access code for using the TPH. Thus, even if the configured hardware token is stolen or compromised, an attacker cannot use the token without knowing the valid hardware access code. In a sense, it achieves security using the paradigm of *"Something you know, and something you get"* providing dual security to the network. This concept is similar to the one used by banks for cash dispensation at ATM's (a combination of card and PIN is required to access the account).

Based upon our experience of using the prototype, we have found that implementation of UGSP can be done in cost-effective way. UGSP is scalable and robust to addition of new members in the User Group. In this paper we have demonstrated and discussed UGSP for data transfer in a User Group in mobile ad-hoc network. However, there are certain generalizations possible as stated follows:

Communication Protocol Independent: Using UGSP, we are able to establish a secure communication channel between nodes at the end of Phase 1. Once this happens, we can use any of the existing protocols, such as TESLA, for data communication.

Multiple Applications: Although, we have chosen data transfer as a sample application for the demonstration of the protocol, UGSP can be used for several purposes like authenticated routing, node-to-node key agreement and ubiquitous computing.

Network Infrastructure Independent: UGSP has been developed for mobile ad-hoc networks, but it is equally efficient in wired networks as well. It replaces PKI in the sense that there is no need to go to the trusted third party everytime you want to validate any certificate.

Membership to Multiple DUG: The TPH token can be made to have more than one location for storing the GAC. A node can thus be a valid user in more than one different ad-hoc networks simultaneously. $iButton$, for example, has eight locations for storing GAC. We are evaluating the system performance when a node is a part of eight simultaneous ad-hoc networks.

REFERENCES

Bobba, R. B., Eschenauer, L., Gligor, V., and Arbaugh, W. A. (2002). Bootstrapping security associations for routing in mobile ad-hoc networks. In *Technical Report, Institute for Systems and Research, UMD, TR 2002-44*.

Hu, Y., Johnson, D., and Perrig, A. (2002). Sead: Secure efficient distance vector routing for mobile wireless ad hoc networks. In *Workshop on Mobile Computing Systems and Applications, IEEE*.

Hu, Y.-C. and Perrig, A. (2002). Ariadne: A secure on-demand routing protocol for ad hoc networks. In *Mobicom*.

Khaili, A. and Arbaugh, W. A. (2002). Security of wireless ad hoc networks. In *http://www.cs.umd.edu/ aram/wireless/survey.pdf*.

Khalili, A. and Arbaugh, W. (2003). Toward secure key distribution in truly ad-hoc networks. In *IEEE Workshop on Security and Assurance in Ad-Hoc Networks*.

Lamport, L. (1981). Password authentication with insecure communication. In *Communications of the ACM, pg. 770-771, Number 81, Volume 24*.

Papadimitratos, P. and Haas, Z. (2002). Secure routing for mobile adhoc networks. In *Communication Networks and Distributed Systems Modeling and Simulation Conference*.

Perrig, A., Canetti, R., Tygar, J., and Song, D. (2002a). The tesla broadcast authentication protocol. In *RSA Cryptobytes*.

Perrig, A., Szewczyk, R., Tygar, J., Wen, V., and Culler, D. E. (2002b). Spins: Security protocols for sensor networks. In *Wireless Network Journal (WINE)*.

Royer, E. M. and Toh, C. K. (1999). A review of current routing protocols for ad hoc mobile wireless networks. In *IEEE Personal communications*.

Stajano, F. and Anderson, R. (1999). The resurrecting duckling: Security issues for ad-hoc wireless networks. In *Proceedings of the 3rd AT & T Software Symposium*.

Toh, C. K. (2001). Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks. In *IEEE Communications Magazine*.

Zhou, L. and Haas, Z. (1999). Securing ad hoc networks. In *IEEE Network Magazine, 13(6)*.

PART 4

Multimedia Signal Processing

IMAGE AUTHENTICATION USING HIERARCHICAL SEMI-FRAGILE WATERMARKS

Yuan-Liang Tang* and Chun-Hung Chen
Department of Information Management
Chaoyang University of Technology
168, Jifong E. Rd., Wufong Township, Taichung County 41349, Taiwan (R.O.C.)
**Email: yltang@cyut.edu.tw*

Keywords: Image authentication, Digital watermarking, Semi-fragile watermarks, Wavelet transformation.

Abstract: In this paper, a semi-fragile watermarking technique operating in the wavelet domain is proposed. A hierarchy of the image blocks is constructed and the image features are extracted such that relationships among image blocks are established in order to enhance the security and robustness of the system. With such a hierarchy, the image can be authenticated at different levels of resolution, hence providing a good property of tamper localization. In addition, by varying certain parameters, the system is able to control the degree of robustness against non-malicious attacks. The proposed algorithm thus provides a fine trade-off between security and localization, and is also robust to common image processing operations.

1 INTRODUCTION[1]

With the advance of network technologies and the popularity of digital multimedia, it is very easy to create, duplicate, transmit, and modify digital products. However, serious problems also arise along with such convenience, that is, unauthorized modification on digital products becomes very easy, too, and detection of such tampering is extremely difficult. If the digital products are images, we face the problem of *image authentication*, namely, to identify if an image has experienced malicious tampering. One of the solutions referred to as *exact authentication* embeds *fragile watermarks* (Lin, 1999) in the image and they break easily even if the image experiences only tiny modification. The applications of exact authentication are very limited because manipulations which preserve the semantics of the image should be acceptable. Such a requirement leads to another solution known as *inexact authentication*, in which *semi-fragile watermarks* (Bartolini, 2001) are embedded in stead of fragile ones. Semi-fragile watermarks are relatively robust to content-preserving manipulations, while fragile to malicious modification.

There are a number of works related to semi-fragile watermarks. For example, Queluz (1999) generated digital signatures, based on moments and edges, to protect the image. An image may be corrupted without affecting their moments, but their edges will certainly be changed. This property is used to authenticate the image content. Yu *et al*. (2000) used the Gaussian distribution to model the amount of modification on wavelet coefficients which is introduced by incidental distortions or malicious attacks. The number of coefficients necessary for watermark embedding is optimized as well. Lin *et al*. (2000) embedded a pseudorandom *m*-sequence into the median frequency DCT coefficients for image authentication. They used correlation values to determine the authenticity of an image.

[1] This research is supported by a grant from National Science Council, Taiwan, R.O.C. (NSC92-2213-E-324-024).

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 241–246.

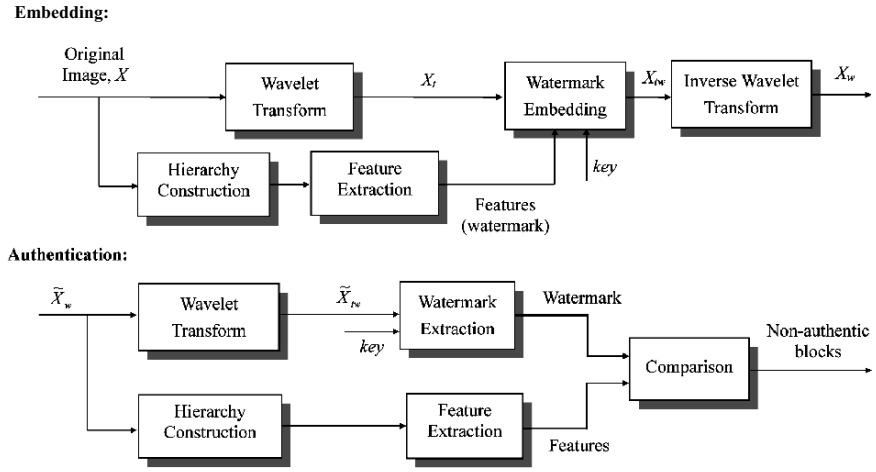


Figure 1: The embedding and authentication processes.

In addition to authenticating an image as a whole, it is also desirable to pinpoint the locations of tampering. Most localized authentication methods rely on some form of block-wise authentication (Wong, 1999), in which the image is divided into disjoint spatial regions and each of them is authenticated independently. The spatial acuity with which a block-based authentication system localizes tampering depends on the block size, thus it might be desirable to reduce the block. However, as indicated by Holliman and Memon (2000), there are potential security risks associated with smaller block sizes—the system is vulnerable to the *collage attack* (Wong, 1999; Yeung, 1997). Therefore, there exists a trade-off between security and localization. In this paper we describe a semi-fragile watermarking technique operating in the wavelet domain. A hierarchy of the image blocks is constructed and the image features are extracted such that the relationships among image blocks are established in order to resist the collage attack. The hierarchy is based on the work done by Celik *et al*. (2002), and with such a hierarchy the image can be authenticated at different levels of resolution, hence providing a good localization property. In addition, the proposed system is designed in such a way that by varying certain parameters, it is possible to control the degree of robustness against non-malicious attacks. Therefore our algorithm provides a fine trade-off between security and localization and is also robust to common image processing operations.

2 THE PROPOSED SYSTEM

Figure 1 delineates the embedding and authentication processes of our system. In the embedding process, a hierarchical structure is first constructed for the original image, X, and then the features are extracted from such a hierarchy. The features represent the image content and are embedded as a watermark into the wavelet-transformed image, X_t, resulting in X_{tw}. Finally, X_{tw} is inversely transformed back to the original format, producing the watermarked image X_w. The embedding process also requires a private key to ensure the security of the system. During authentication, both the image features and the watermark are extracted using the same methods as in the embedding process. These two pieces of data are then compared against each other to determine if the image blocks are authentic. The locations of tampering, if any, will be reported as well.

2.1 Construction of the Hierarchy and Feature Extraction

Celik *et al*. (2002) proposed a hierarchical structure in which the original image is first divided into disjoint blocks which constitute the bottom level of the hierarchy. And then successive levels are formed by combining distinct groups of blocks at a preceding level. Without loss of generality, we assume that 8-bit grayscale square images are dealt with. Given an $N \times N$ image X, if X is divided into

$M \times M$ blocks at the bottom level, we have a hierarchy of $L = \log_2 M + 1$ levels. Let X_{ij}^l denote a block at level l, $l = 0..L-1$, where indices ij represent the spatial position. Assuming that 2×2 blocks at a given level of the hierarchy are combined to create a block at the next level, we have

$$X_{ij}^l = \begin{bmatrix} X_{2i,2j}^{l-1} \,||\, X_{2i,2j+1}^{l-1} \\ X_{2i+1,2j}^{l-1} \,||\, X_{2i+1,2j+1}^{l-1} \end{bmatrix},$$

for $l = 1..L-1$. The top level thus consists of only one block $X_{00}^{L-1} = X$. Based on Celik's hierarchy, for each block, X_{ij}^l, we first compute the mean, m_{ij}^l, of pixel intensities of the block. Due to the limitation of capacity, the bottom-level mean, m_{ij}^0, is quantized into 64 levels, i.e., a 6-bit intensity instead of the ordinary 8-bit intensity. In addition, since tampering with a block may not affect the mean when the block size is large, we introduce the *polarity* to improve the sensitivity as well as the reliability of detection. The four-bit polarity, p_{ij}^l, of X_{ij}^l is obtained by comparing the parent block's mean with those of its 4 children:

$$p_{ij}^l(x,y) = \begin{cases} 1, \text{ if } m_{ij}^l \geq m_{2i+x,2j+y}^{l-1} \\ 0, \text{otherwise} \end{cases},$$

for $l = 1..L-1$ and $x, y = 0..1$. Denoting $|\cdot|$ as the length in bits, we have $|m_{ij}^0| = 6$, $|m_{ij}^l| = 8$, and $|p_{ij}^l| = 4$ ($l = 1..L-1$), respectively. These intensity means and polarities, denoted by A_{ij}^l ($A_{ij}^0 = m_{ij}^0$ and $A_{ij}^l = m_{ij}^l \,||\, p_{ij}^l$, $l = 1..L-1$), are regarded as the image features (i.e., the authentication data or watermark) and are embedded back into the image for content protection.

2.2 Watermark Embedding

The coefficients in frequency band LL_2 of the wavelet-transformed image are selected for embedding. These coefficients are good candidates in that they represent the perceptually significant part of the image and it is impossible for an attacker to tamper with the image without gross modifications to its appearance. The high level authentication data is spread over a number of lower level blocks and the accumulated payload is inserted at the lowest level of the hierarchy by wavelet coefficient modification. This is done by partitioning A_{ij}^l into a number of smaller strings:

$$A_{ij}^l = A_{ij}^l\{0,0\} \,||\, A_{ij}^l\{0,1\} \,||\, ... \,||\, A_{ij}^l\{\Lambda(l)-1, \Lambda(l)-1\},$$

where $\Lambda(l) = 2^l$. The payload of a block on the lowest level is formed by concatenating the units inherited from higher level blocks:

$$D_{ij} = A_{ij}^0 \,||\, A_{C_1(i),C_1(j)}\{i - C_1(i), j - C_1(j)\} \,||\, ... \,||\, A_{C_{L-1}(i),C_{L-1}(j)}\{i - C_{L-1}(i), j - C_{L-1}(j)\},$$

where $C_b(x) = \lfloor x/2^b \rfloor$. After the above preparation, wavelet coefficients corresponding to each block on the lowest level of the hierarchy are embedded with payload bits. To increase the security level of the system, we use the pseudo-random number generator (PRNG), initialized by a private key, to establish the correspondence between an image block and the wavelet coefficients. This is illustrated in Figure 2, in which the watermark is embedded in the corresponding 4×4 coefficients in subband LL_2.

Kundur and Hatzinakos (1999) embed the watermark by first defining the quantization function:

$$Q(f,q) = \begin{cases} 1, \text{ if } kq \leq f < (k+1)q \text{ for } k = 0, \pm 2, \pm 4 ... \\ 0, \text{ if } kq \leq f < (k+1)q \text{ for } k = \pm 1, \pm 3, \pm 5 ... \end{cases}$$

where f is the wavelet coefficient and q denotes the size of the quantization interval. They update f by

$$f' = \begin{cases} \Delta f + 0.5q, \text{ if } Q(f,q) = b \\ \Delta f + 1.5q, \text{ if } Q(f,q) \neq b \text{ and } r > 0.5q \\ \Delta f - 0.5q, \text{ if } Q(f,q) \neq b \text{ and } r \leq 0.5q \end{cases}$$

where $\Delta f = \lfloor f/q \rfloor \cdot q$, $r = f - \Delta f$ (quantization noise), and b is the watermark bit. Obviously, the result of such update will locate at exactly the middle of the quantization step, which makes it very easy to identify the watermarked coefficients. To overcome this security risk, we modify the coefficient update function as follows:

$$f' = \begin{cases} f, \text{ if } Q(f,q) = b \text{ and } (0.5q - z) \leq r \leq (0.5q + z) \\ \Delta f + 0.5q + s, \text{ if } Q(f,q) = b \text{ and } r > (0.5q + z) \\ \Delta f + 0.5q - s, \text{ if } Q(f,q) = b \text{ and } r < (0.5q - z) \\ \Delta f + 1.5q - s, \text{ if } Q(f,q) \neq b \text{ and } r > 0.5q \\ \Delta f - 0.5q + s, \text{ if } Q(f,q) \neq b \text{ and } r \leq 0.5q \end{cases}$$

where s is a random number in the range $[1..z]$ and z is the randomness tuner ($z = \lfloor q/6 \rfloor$ in our experiments). The result of such new update will look random and therefore is more secure. Normally, a larger q gives a more robust watermark and it should vary according to the host image. However, a larger q also creates more visual impact. In order to search for an appropriate value, dozens of well-known images were tested to obtain the relationship between q and

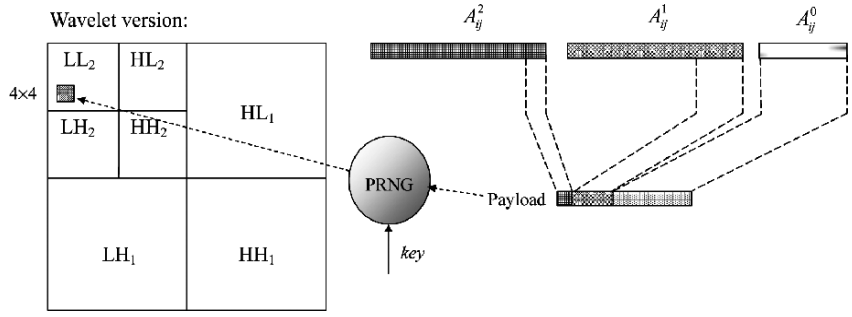


Figure 2: Concatenation of blocks to form a payload and placement of resulting payload in wavelet domain of the image.

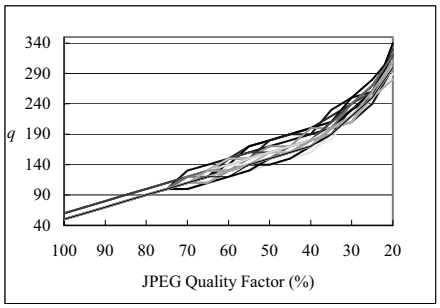


Figure 3: Watermark robustness as a relationship between quantization intervals and compression quality factors.

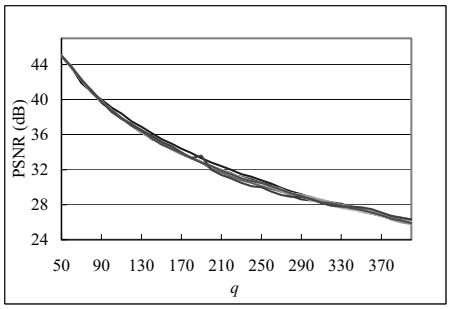


Figure 4: PSNR versus quantization intervals.

the JPEG quality factors (*QF*s). Figure 3 shows the results, in which, for a specified value of *QF*, any quantization intervals set below the curve will cause our system to produce false positives. For instance, if we want to validate the watermarked image that can withstand JPEG compression of 80% *QF*, *q* should be greater than 85, otherwise such an operation will be identified as a malicious attack instead of common processing. In order to obtain a more general form, those curves are approximated by a second order non-linear equation using the regression technique, and the following equation is obtained:

$$q = 393 - 6.014QF + 0.027QF^2.$$

The equation above actually defines the lower bound of the quantization interval. On the other hand, to determine upper bound, the PSNRs as a function of *q* are computed for each watermarked image. Figure 4 shows the results and, because PSNR ≥ 30 is generally acceptable, *q* should be less than 200, namely, the upper bound. Since the lower bound produces least visual impact on an image, it is a good candidate when determining *q*. As a consequence, our system allows determining the quantization interval automatically, depending on the visual quality requirement of JPEG compression.

2.3 Image Authentication

The authentication process is analogous to the embedding process. Let $\hat{X}_{ij}^l$ denote a block of the image that may have been tampered with. The watermark bit is extracted by $\hat{b} = Q(\hat{f}, q)$, where $\hat{f}$ is the corresponding coefficient. The partitioning algorithm used during embedding is reversed to recover the authentication data $\hat{A}_{ij}^l$, which is further partitioned to obtain $\hat{m}_{ij}^l$ and $\hat{p}_{ij}^l$. The same feature

extraction is also applied to obtain $\tilde{m}_{ij}^l$ and $\tilde{p}_{ij}^l$ for each block. And finally, the difference between the extracted features and watermark is calculated. Let $T_{ij}^l = \left|\tilde{m}_{ij}^l - \hat{m}_{ij}^l\right|$; $\hat{X}_{ij}^l$ is determined as non-authentic if $T_{ij}^l > T^l$, where T^l is the threshold and it varies according to the size of the block. At the bottom level, since we have ignored the 2 least significant bits when collecting the authentication data, we set T^0 = 8 (3 bits) to increase the robustness. Furthermore, because tampering with a small area may have little influence on the intensity mean of a large-sized block, the threshold should be smaller. In our experiments, we set T^1 = 6, T^2 = 4, and T^l = 2 for l = 3..L–1 to accommodate such a situation.

For polarity checks, $\hat{p}_{ij}^l$ and $\tilde{p}_{ij}^l$ are compared against each other bit by bit. Any bit difference signifies a non-authentic block. However, if the intensity means of the two blocks are similar, non-malicious modification may easily reverse their polarity. Based on such reasoning, when the intensity difference between the parent block and the child block is small, say less than 4, that bit is ignored during comparison. In summary, a block is authentic only when it passes both intensity mean and polarity tests.

3 EXPERIMENTAL RESULTS

In our experiments, the 512×512 grayscale *Lena* image is used as the host image, as shown in Figure 5(a). We set the size of the lowest level block to be 16×16 pixels, which results in a 6-level hierarchy. Figure 5(b) shows the watermarked image, whose PSNR value is about 38 dB. The degradation of the watermarked image depends on the amount of the embedded data and the embedding strength. To demonstrate the effectiveness of our technique, we modify the image by placing a tattoo (apple) on *Lena*'s arm (Figure 5(c)). As can be seen in Figure 5(d), the tampered blocks are correctly detected, in which non-authentic blocks at lower levels are shown in darker shades, while those at upper levels are shown in lighter shades. Furthermore, we perform several non-malicious manipulations to test the robustness of our system, including 80%-*QF* JPEG compression, blurring, sharpening, and addition of Gaussian noise with zero mean and variance of 20. As expected, our system didn't make any false positive errors and Table 1 shows the results.

4 CONCLUSION

We have presented in this paper an image authentication technique using semi-fragile watermarks. The authentication data is embedded in the image and is arranged in a hierarchical structure so that the whole contents of image are tightly connected in order to overcome the security weakness of block-based techniques. The system is insensitive to common image processing techniques in that robust image features are selected and a variable quantization interval further controls the degree of robustness. The system is also secure because not only the block-dependence property significantly discourages the collage attack, but also the random correspondence between blocks and coefficients prohibits brute-force attacks. The experimental results demonstrated that our system is very effective.

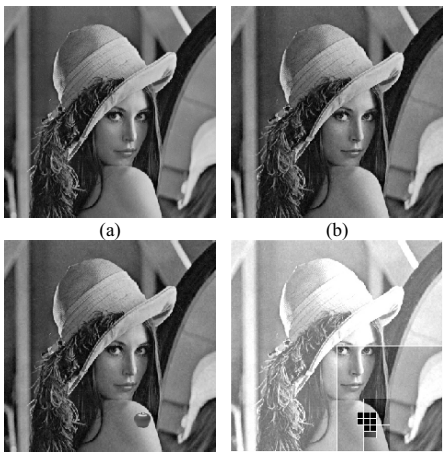


Figure 5: (a) Original image, (b) watermarked (q = 110), (c) tampered, (d) detection result.

Table 1: Experimental results of various attacks.

Attack	Quantization Interval	Authentic?
JPEG (QF = 20%~100%)	By formula	Yes
Blurring	160	Yes
Sharpening	190	Yes
Gaussian noise addition	120	Yes

REFERENCES

Bartolini, F., Tefas, A., Barni, M., and Pitas, I. (2001) "Image authentication techniques for surveillance applications," *Proc. IEEE*, vol. 89, no. 10, pp. 1403–1418.

Bhattacharjee, S. and Kutter, M. (1998) "Compression tolerant image authentication," *Int. Conf. Image Processing*, vol. 1, pp. 435–439.

Celik, M.U., Sharma, G., Saber, E., and Tekalp, A.M. (2002), "Hierarchical watermarking for secure image authentication with localization," *IEEE Trans. Image Processing*, vol. 11, no. 6, pp. 585–595.

Chotikakamthorn, N. and Sangiamkun, W. (2001) "Digital watermarking technique for image authentication by neighbouring block similarity measure," *Int. Conf. Electrical and Electronic Technology*, vol. 2, pp. 743–747.

El-Din, S.N. and Moniri, M. (2002) "Fragile and semi-fragile image authentication based on image self-similarity," *Int. Conf. Image Processing*, vol. 2, pp. 897–900.

Fridrich, J., Goljan, M., and Du, R. (2001) "Invertible authentication watermark for JPEG images," *Int. Conf. Coding and Computing*, pp. 223–227.

Holliman, M. and Memon, N. (2000) "Counterfeiting attacks on oblivious block-wise independent invisible watermarking schemes," *IEEE Trans. Image Processing*, vol. 9, pp. 432–441.

Kundur, D. and Hatzinakos, D. (1999) "Digital watermarking for telltale tamper proofing and authentication," *Proceedings of the IEEE*, vol. 87, pp. 1167–1180.

Lin, E.T. and Delp, E.J. (1999) "A review of fragile image watermarks," *Multimedia and Security Workshop* (*ACM Multimedia* '99) *Multimedia Contents*, pp. 25–29.

Lin, E.T., Podilchuk, C.I., and Delp, E.J. (2000) "Detection of image alteration using semi-fragile watermarks. *SPIE Conf. Security and Watermarking of Multimedia Content II*, vol. 3971, pp. 152–163.

Queluz, M.P. (1999) "Content-based integrity protection of digital images," *SPIE Conf. Security and Watermarking of Multimedia Contents*, vol. 3657, pp. 88–93.

Sun, Q. and Chang, S.-F. (2002) "Semi-fragile image authentication using generic wavelet domain features and ECC," *Int. Conf. Image Processing*, vol. 2, pp. 901–904.

Wong, P.W. (1999) "A watermark for image integrity and ownership verification," *IS&T Image Processing, Image Quality, Image Capture, Systems Conference*, pp. 374–379.

Yeung, M. and Mintzer, F. (1997) "An invisible watermarking technique for image verification," *Int. Conf. Image Processing*, vol. 2, pp. 680–683.

Yu, G.J., Lu, C.S., Liao, H.Y. Mark, and Sheu, J.P. (2000) "Mean quantization blind watermarking for image authentication," *Int. Conf. Image Processing*, vol. 3, pp. 706–709.

DEPLOYMENT OF LIVE-VIDEO SERVICES BASED ON STREAMING TECHNOLOGY OVER AN HFC NETWORK

David Melendi[1], Xabiel G. Pañeda[1], Roberto García[1], Ricardo Bonis, Víctor G. García[1]
Computer Science Department, University of Oviedo
Campus Universitario de Viesques. Sede Departamental Oeste, 33204 Xixón-Gijón, Asturies
{melendi, xabiel, victor, roberto}@correo.uniovi.es[1]

Keywords: Live, Video, Streaming, Multimedia, HFC.

Abstract: This paper presents an approach to the deployment of a live-video service based on streaming technology over an HFC network. This approach covers most of the issues that may arise while putting one of these services into operation, taking into account new aspects such as those oriented to the improvement and prior analysis of the service's behaviour. An accurate and continuous service analysis can contribute to boost the service's performance and thus to lead the service to the so called *excellence of service*. This paper also presents a service architecture specifically designed for HFC networks that takes advantage of the structure of this kind of networks. Furthermore, a complete framework that facilitates most of the tasks that are needed to deploy and manage a live-video service over the internet is presented.

1 INTRODUCTION

The emergence of the World Wide Web has changed the Internet world. This service has become a powerful medium. Daily, an important number of web accesses is produced and a huge volume of information is delivered. The bandwidth increase in subscribers' access capabilities has given rise to the appearance of a new complementary service: the Internet video. There are two types of video services on the Internet: live-video and video-on-demand. In video-on-demand services, the user requests the information at any time and the server delivers it exclusively. This system allows users to interact with information: Pauses, backward and forward jumps are allowed. Its behaviour is similar to a videotape. On the other hand, in live-video services, contents are received directly by the server, which broadcasts them straight out to the audience.

Nowadays, most video services on the Internet are based on streaming technology. The advantages of video streaming and the subscribers' expectations are important. However, this technology presents some problems. Video delivering consumes an important bandwidth in the network and requires a constant quality of service. What is more, live-video services require much more transmission capabilities than video-on-demand services, due to the fact that all the users connect at the same time. To maintain service quality under control and select the most interesting contents, the use of proper engineering techniques and good analysis methods is fundamental. The analysis systems must provide the necessary information to ensure the correct configuration of the streaming service, and take as much advantage as possible of the subjacent network technology.

In this paper, an approach to engineering and analysis methods for live-video services over HFC networks is presented. The main aim of this work is to provide useful tips to help service managers in planning, deploying, configuring and improving live-video services. Furthermore, the paper has followed an interesting practical approach, based on the improvement of these services through the analysis of the information provided by existing technologies.

The improvement in the transmission of multimedia contents over the internet is a fact in the current research world. There are abundant papers that cover most of the topics related to the technologies involved in the distribution of live-video contents. Some of them, such as (Chow, 2000) or (Turletti, 1994) commented on new engineering techniques to deploy live-video services, but assuming the availability of multicast technologies. Others like (Ortega, 2000) or (Tham, 2003), are mainly oriented to the study or the development of new data formats for the transmission of live-video. There are others such as (Chawathe, 2000),

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 247–254.

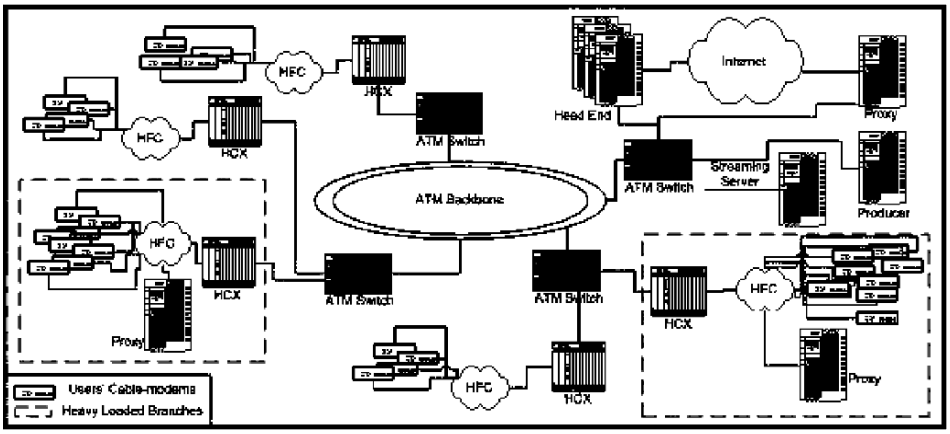


Figure 1: Service Architecture.

(Deshpande, 2001), (Nguyen, 2002) or (Padmanabhan, 2002), that offer different approaches to the deployment of a streaming service over a network, but using proprietary solutions or basing their research on service models or simulations.

Although some of the topics covered in this paper have been revised in other publications, the main difference is the practical point of view that has been followed. The conclusions have been obtained through the analysis of the available data from one of these services and the solutions have been designed to improve a real service.

The rest of the paper is organized as follows: Section 2 shows the proposed service architecture over an HFC network. A detailed explanation of live-video services engineering methods is set out in section 3. Section 4 covers an approach to live-video services analysis. An introduction to service configuration is offered in Section 5. Finally, conclusions will be presented in section 6.

2 SERVICE ARCHITECTURE

A live-video service requires the installation of several devices to support the content distribution over the network. The main components that any of these services need are the production software, the streaming server, a set of proxies and the multimedia clients that should be installed in the customers' computers. The distribution of these devices over the network is clearly connected with the type of networking technology that is being used, and the future performance of the service will be determined by the placement of each of the systems involved. Figure 1 shows the proposed distribution over an HFC network.

HFC networks are commonly structured hierarchically around a central spot that delivers all the services to the users that are connected to the network (García, 2003). Although the physical structure of most of these networks may seem different due to the use of ATM backbone rings or other redundant architectures, the logical structure is always hierarchical around this central point, called the *head end*. The *head end* manages all the services in a centralized way: the accesses to the internet, the compilation and distribution of the TV channels, the connection to the telephone networks, etc. It is also in charge of assigning the proper resources to the users whenever they try to use one of the services provided. Therefore, the best place to install the streaming server, whose mission is to deliver contents to the users, is precisely close to the *head end*. This location will permit both a better management by the owner of the network and an increased assignment of output bandwidth rate, in order to avoid problems while distributing the contents to the network.

On the other hand, the mission of the production software is to capture live or stored contents, adapt them for streaming transmission, and deliver them to the streaming server. This device should be as close to the streaming server as possible, in order to avoid cuts during the transmission of contents between both systems. If the contents are being captured live in a remote location with access to the HFC network, a proper *constant bit-rate* connection should be allocated to preserve transmission quality. If that location is outside the HFC network, two alternatives need to be considered: either to subcontract an external connection, or to store the

contents and retransmit the saved files later. If an external connection is subcontracted, the external provider must guarantee transmission quality on the route between the producer and the streaming server.

The optimal type of connection for live transmissions is a multicast connection, which reduces the amount of traffic in the network and the load on the server. But multicasting can not be used in most of the existing networks due to hardware incompatibilities, so proxies could be used in order to improve network performance. The mission of proxies is to receive multimedia streams from the main server and retransmit them to final customers or to other proxies. Furthermore, the use of this kind of devices may reduce the load on the streaming server, and avoid possible cuts during the transmission of contents due to a hypothetical overload of that machine. Proxies can also be installed following an *on-cascade* service architecture. This architecture allows proxies to serve contents to other proxies, acting as servers, and reduces the load on the main server.

Every heavy-loaded branch of the HFC network should have, at least, one proxy running in order to serve the customers in that branch. If one proxy is not enough to serve one of those branches, more can also be placed following the *on-cascade* architecture mentioned before. On the other hand, branches with a small number of users can be served from a remote proxy, possibly allocated at the *head end* with proper connection capabilities. If the contents are also going to be delivered outside the HFC network, an additional proxy could be placed to attend all the requests coming from the Internet.

Should the service be offered to external connections, it is also important to consider the placement of several proxies in the networks used by potential users, through some kind of service level agreements with the corresponding access providers.

3 SERVICE ENGINEERING

The deployment of any high-cost service that may suffer problems due to several different circumstances, requires an intense development of engineering tasks in order to reduce service costs, improve service performance and increase customer satisfaction. These engineering tasks should be oriented to improve the service in the following areas: the network, devices and contents.

The network is a critical aspect in any distributed service. It is even more critical in services like live-video distribution, where contents need to be sent with a constant rate to avoid cuts during their reproduction in the customers' computers. Although the optimal network design for these services is not always available, the use of some alternative solutions may mitigate most of the transmission problems that can arise during the delivery of contents. In most cases, transmission difficulties appear in the network's segment known as *last-mile*. One of the features of HFC technology is that it combines optical fibre and coaxial cable infra-structures, relegating the latter to the last extreme of the network, shared between 100 and 200 customers. The fact that these network extremes work under a best effort strategy, combined with the limits of the coaxial cable, reduces transmission capabilities and the network's grade of scalability. If there are a high number of users that demand the transmission of live-video contents in one of these extremes, the only way to avoid transmission problems is to bring the optical fibre closer to users, or to reduce the number of users that can be connected to the network in those extremes. It is clear that these solutions are not always feasible, so the only way to deliver live contents to those users is to produce them with a decreased video quality.

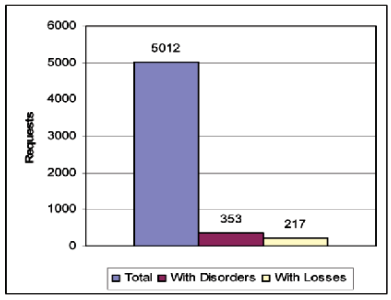


Figure 2: Requests with delivery problems.

There are also technologies available in the market designed to ensure the content delivery, such as *surestream* (RealNetworks, 2002). This technique is capable of adapting contents' quality in real time, depending on the transmission capacity that is perceived in the customers' computers.

To detect transmission problems it is necessary to analyze the network's behaviour, and both the server and proxies log files in order to identify late arrival of packets, disorder of packets, loss of packets, reduced reproduction times, etc. Figure 2 shows requests with delivery incidences registered during a real live event.

Although not common, there are sometimes other problems produced in the network due to incorrect routing configurations that may produce the loss of packets or their late arrival. The existence of this kind of problems may affect not only the transmission quality of live-video, but also the

quality of all the delivered data. Although the detection of this type of errors is even harder than in the previous case, a high loss packet rate or late packet rate of the customers of a determined network branch, can be the definitive clue to identify incorrect routing policies in the network. Again, the solution to this problem can be found through the analysis of the server and the proxies log files.

Moving forward, the second issue that was expressed as critical was related to the devices that are being used during the transmission of live contents. The simpler live-video service consists of a machine where both the production and the distribution software are running. This initial configuration may suffer several problems such as a high CPU load, huge memory consumption, elevated hard disk utilization and a possible overload in the output connection to the network.

The execution of both programmes in the same machine may overload the computing capacity of the latter, and so affect service performance in a severe way. It must be taken into account that as users' requests reach the server, higher resources are needed to maintain service quality. It is necessary to observe CPU load and memory consumption in order to detect performance problems in this kind of services. If overload errors occur, an inexpensive investment is to dedicate one machine to produce the contents and another to host the streaming server.

This new configuration requires a high connection quality between both devices. If a direct or dedicated connection is not possible, it is essential to analyse the producer's log files to detect problems that may arise during the delivery of contents.

It is necessary to comment that there are some connection policies used in commercial applications that do not report about transmission problems between the producer and the server. An example is one of the *push* methods provided by *Realnetworks' Helix Producer*, where streaming servers do not establish a feedback channel with the producers. Special care must be taken in these cases, and other connection methods should be used if quality can not be assured. As far as the connection method is concerned, this will depend on the distribution and the number of connections that the server receives. If there is a constant connection rate in the server, one of the available *push* methods should be used. On the other hand, if there is a variable arrival of requests, a *pull* connection may be the best solution to save resources in both machines.

Although the split of production and delivery applications between two computers is a clear improvement, a high connection rate in the server may cause the previously commented overload. If all the requests are attended by a single machine, several problems may again be encountered: high CPU utilization, memory overload, elevated bandwidth consumption, and license limitations.

Commercial licenses usually affect the number of simultaneous connections, or the output bandwidth that servers can handle. If delivery problems are being caused by license restrictions, the simplest solution is to acquire a less restrictive license. To detect this type of problems, it is necessary to observe the server's log files, calculate all the simultaneous connections that are being handled in every moment, and compare them to the number of simultaneous connections that are permitted by the existing license. It is also necessary to calculate the output bandwidth that is being used, and compare it with both the license limitations and the capacity of the line that is being used to deliver the contents to the users. If there is high bandwidth consumption in the server's output, network reengineering must be carried out in order to mitigate these problems. More capacity should be allocated, or clustering solutions should be applied by distributing several proxies in the network that will support the delivery of contents to the users. The latter solution is also applicable when performance problems have been detected in the machine that hosts the streaming server, and a computing capacity increase is not feasible.

Proxies are in charge of forwarding the contents to the users. Although in *on-demand* transmissions they operate following caching strategies, in live transmissions they mainly receive the streams sent by other devices and forward them to the users that request the contents. The origin devices could be the main server or another proxy that works under an *on-cascade* architecture.

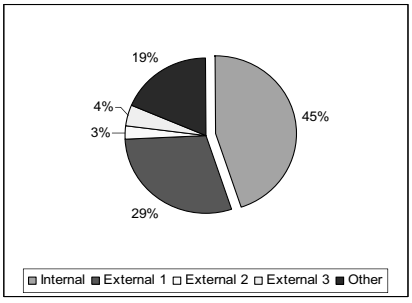


Figure 3: Origin of Requests.

In networks where multicasting is not available, proxies can be used to bring the transmission closer to users, reducing the load on the main server and decreasing traffic in the network. In HFC networks, proxies could be allocated in heavy-loaded branches where there is an important number of users requesting the transmission of contents. A step forward

is to install several proxies *on-cascade*, depending on the evolution of demand in those branches, or in the load that has been registered in the proxies. On the other hand, network branches with a low number of requests could be served directly from the main server, or for further performance, from a centralized proxy that could be used to redirect transmissions to external networks. In any case, it is very important to collect data from the network and the proxies that have been installed, and analyze said data in order to detect possible performance deficiencies or loss in transmissions. Figure 3 shows the origin of users' requests, registered during a real live event.

An extremely important issue is that of *content management*. Contents are usually provided by a different entity than the network operator. Sometimes it is a communications media, such as a TV company or a digital newspaper, other times a movie producer, and most of the times a media management company that sells contents to other businesses. Once those contents are delivered through the network, it is very important to analyze whether they have been successful or not. An inadequate selection of contents may greatly influence the budget of the service or its profitability. Although it is very difficult to calculate audience statistics in other services such as conventional TV, with live-video transmissions it is possible to obtain detailed information about users' accesses. There are different aspects that could drive the production of contents, and are available in this kind of services: number and length of connections, preferred time ranges, users' installed language and computing capacity, etc. These are very important data that should not be underestimated. Servers and proxies log files provide this type of information that needs to be analyzed in detail in order to calculate user's satisfaction and preferences. This information is usually owned by network operators, who could give consultancy support, or reporting services to content providers.

4 APPROACH TO SERVICE ANALYSIS

Once the service has been deployed over the network, it is necessary to monitor the transmissions and check if everything is working properly. It must be taken into account that live-video services do not allow second chances, after they occur their live transmission is no more interesting. Other services such as video-on-demand could be improved using continuous analysis and configuration cycles, but live-videos are slightly different due to their temporary nature. Errors during a live transmission are complete failures, so everything must work properly to ensure the success of the service.

Although live-video transmissions with problems can not be fixed, their analysis can be considered as a continuous learning tool to improve future emissions. The traditional *learn through experience* thesis is perfectly applicable to these services. So it is necessary to analyze live-video transmissions to know what is happening, why it is happening and how it can be improved.

The analysis of live-video services consists of the detailed observation of three of the different stages that can be identified in any live transmission: production, distribution and visualization. Hence the division of service analysis in the following phases: Production Analysis, Distribution Analysis and Visualization Analysis. At the same time, these three analyses consider the issues that were laid down in the previous section – network, devices and contents- from different points of view.

4.1 Production Analysis

Production analysis is centred on the contents production phase. During this stage, the contents are captured and coded using a particular algorithm. After digitalizing contents in the proper format, they are sent to the server using the streaming technology. It is necessary to ensure that the device that is in charge of this task does not suffer any performance incidence. It is also very important to check the connection between the producer and the streaming server. Among others, such as CPU throughput, or memory consumption, the following quality metrics can be used for the analysis of production phases: Production Loss Rate and Production Bandwidth Consumption.

Production Loss Rate calculates losses in the transmission channel between the producer and the streaming server. It can be obtained through equation 1.

$$PLR = \frac{RP - SR}{PS} \tag{1}$$

Where RP is the number of resent packets, SR the amount of successful resends and PS the number of packets sent to the streaming server. All this data can be gathered from the producer's log files.

This metric is designed to calculate the losses of information during the production phase, generated by problems in the connection between the production software and the streaming server. It must be taken into account that, although some transmission problems can be mitigated thanks to the input buffer allocated in the streaming server, severe conditions in the connection between both devices

can mean an important decrease in the quality of the service. Should these problems appear, an improvement in the network infrastructures needs to be requested in order to guarantee a constant transmission quality to assure the delivery of contents to the streaming server.

Production Bandwidth Consumption calculates the bandwidth that live production is consuming. It can be obtained using equation number 2.

$$PBC = \frac{TBR}{AVB} \quad (2)$$

Where *TBR* is the total bit-rate generated in the production phase, and *AVB* is the available bandwidth in the output of the production device. The first parameter is obtained from the producer's log files, adding the output quality that is being generated for each of the targeted audiences of the service, whereas the latter is the bandwidth that is available in the connection where the production device has been plugged into. It is obvious that *AVB* can never be less than *TBR*, because this situation would lead to an increase of the losses in the channel between the production device and the streaming server. Moreover, it must be taken into account that other applications running in the production device may consume output bandwidth, so *PBC* should never be greater than 0.75.

On the other hand, there is no available information in this phase to analyze contents. But it must be taken into account that the media selection is closely related to the analysis of the users' preferences. So this phase depends entirely on the results obtained in the Visualization Analysis phase.

4.2 Distribution Analysis

Distribution analysis is designed to control the quality of the transmissions established between the main streaming server, the proxies and the final customers of the service.

Each of the devices that need to be used to deploy a live-video service over an HFC network, need to be analyzed in detail, to detect performance issues that may affect the final results of the transmission. Hence, it is necessary to analyze the evolution in the resources' consumption of those devices: CPU utilization, memory load, bandwidth consumption, etc. These devices are usually owned by network operators, so no transmission limitations have been considered, except those inherent to the HFC technology and the available network infrastructures. Apart from the typical performance analyses, it is also necessary to consider the license consumption in the main streaming server and the proxies spread throughout the network.

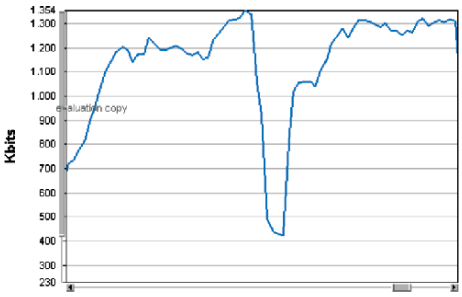


Figure 4: Evolution of *TBR*..

These licenses usually limit the number of concurrent connections accepted by each device, or the output bandwidth that is being dedicated to deliver multimedia contents. It is important to mention this feature, because it can severely damage the growth of the service, rejecting connections requested by new users. The licenses utilization can be obtained calculating one of the equations 3 or 4.

$$ULC = \frac{CC}{MAC} \quad TLC = \frac{TBR}{MBR} \quad (3) \text{ and } (4)$$

Where *ULC* is the users' license consumption, *CC* is the number of current connections, *MAC* is the maximum accepted connections, *TLC* is the transmission's license consumption, *TBR* is the total bit-rate used to deliver the contents, and *MBR* is the maximum bit-rate accepted. If *ULC* or *TLC* reach 1 during long periods of time, it is necessary to consider the acquisition of a higher license. Figure 4 shows the evolution of *TBR* in the output of a streaming server, during a real live event.

It is also necessary to evaluate the origin of requests, in order to detect network branches that may be overloaded due to an elevated number of users, or high network utilization by means of distinct applications like *p2p* clients or other heavy consuming software. As has been said, heavy loaded branches in HFC networks may require the existence of a proxy that could bring the transmission of contents closer users. For these cases, it is good practise to assign specific IP ranges to each of the network branches, to identify the origin of users' requests. This policy may also be useful to locate other transmission problems and solve them with high efficiency and precision.

Another interesting study is to analyze the deterioration of the expected quality, understood as the problems that users are suffering due to an incorrect selection of audiences –or qualities- during the production phase. During the configuration of production, the most critical step is the selection of the audiences that will be supported during the

transmission. If this selection is incorrect, customers may suffer visualization problems due to poor bandwidth availability. The detection of this kind of situations can be done using equation 5.

$$EQD = \begin{cases} OB \geq EB & 0 \\ OB < EB & 1 - \dfrac{OB}{EB} \end{cases} \quad (5)$$

Where *EQD* is the expected quality deterioration, *OB* is the user's obtained bandwidth, and *EB* is the expected bandwidth set during the production phase. The higher this value is, the poorer the reproduction quality has been. An elevated number of high values in this metric should be interpreted as an incorrect selection of audiences during the production phase that needs to be reconsidered for future events.

4.3 Visualization Analysis

Visualization analysis has been designed to check service performance from the users' point of view. Therefore, this analysis considers both the quality of visualization, and the quality of the contents that are being delivered.

Issues regarding quality of visualization are most frequently caused by transmission problems, but users are not aware of the problems that may arise during the delivery of contents. What users are aware of is that sometimes the transmission cuts, the image stops or the initial load time is very high. To bring this analysis closer to users' minds or expectations, all these problems have been grouped into what can be called *Transparency of Service*.

Apart from technology evolution, the different technical solutions or their applicability, the new services that they offer, etc. every single distributed service has one goal, and that is *Transparency*.

When software began to be distributed new problems arose that had not been considered: transmission problems, synchronism issues, format incompatibilities, etc. Live-video, like any other distributed service, has to assure *Transparency*. Users must perceive the reproductions as local to their computers and have to be unaware of the real location of the source of the transmission.

Every incidence that takes place in the delivery of contents, from the production phase to the visualization of the media in the users' computers, has a certain impact on the final reproductions. This impact is a clear deterioration in the *Transparency of Service*. Users' are aware that there is a problem and realize that contents are not stored in their computers. Moreover, they automatically tend to think that this new –or different- product is worse than the previous service they already know, e.g., live Internet video versus conventional TV or video-on-demand. A metric has been developed to evaluate this *Transparency of Service*, using equation 6.

$$ToS = \frac{\lambda * (AQ + VQ + CI) + ES + WC}{3 * \lambda + 2} \quad (6)$$

Where *ToS* is the *Transparency of Service*, *AQ* is the audio quality, *VQ* is the video quality, *CI* is the coefficient of interruption, *ES* is the value of the expected stop metric, *WC* is the waiting coefficient, and λ is the coefficient that adjusts the results of the metric to the preferences of service managers. A value for λ greater than 1 corresponds to analyses that give more importance to the quality of visualization. On the other hand, a value less than 1 gives more importance to the rest of the features.

Audio quality, or *AQ*, is calculated as the percentage of requests without lost or delayed audio packets, and no failed audio resends. Video quality, or *VQ*, is obtained equally to *AQ*, but using video packets information.

On the other hand, the coefficient of interruption, or *CI*, indicates the quality of reproductions from the point of view of buffer reloads. Whenever a client's buffer is consumed, the current reproduction is stopped until new packets have filled a certain amount of this buffer. A high percentage of buffer reloads is symptom of a poor quality in the reproductions. Thus, this coefficient tries to obtain the impact of those interruptions by calculating the percentage of reproductions with no buffer reloads.

The expected stop metric or *ES*, considers the fact that, sometimes, the reproductions do not end for natural reasons, but for transmission problems. Therefore, it tries to estimate the control level that users have while viewing the contents, obtaining the percentage of requests that end with the interaction *STOP*, or because the transmission has finished.

The waiting coefficient, or *WC*, estimates the effects of the time that users have to wait until their reproductions start. During this interval, the communication between the clients and the server is established, and the client's buffer is loaded. If these tasks require too much time, users may feel disappointed and decide to abandon their requests. This metric tries to obtain the influence of this effect by calculating the value of equation 7.

$$WC = \begin{cases} t_{PreRoll} \geq t_{load} & 100 \\ t_{PreRoll} < t_{load} & 100 * \dfrac{t_{PreRoll}}{t_{load}} \end{cases} \quad (7)$$

Where $t_{PreRoll}$ is the estimated load time during the production of contents and t_{load} is the real load time measured in the users' clients.

Once the quality of the reproductions has been checked, it is also very important to ensure that the offered contents meet the customers' preferences.

Several metrics have been developed regarding this issue, the most important being the Impact of the Service or *IoS*. It must be taken into account that while in Web services the only metric that evaluates the quality of contents is the number of accesses, in video transmission two different aspects must be considered: the number of accesses and their length, the information being continuous. *IoS* evaluates both aspects, and checks the quality of the offered contents using equation 8.

$$IoS = \sum \frac{VP * RIU}{100 * IU} \quad (8)$$

Where *VP* is the visualized percentage, *IU* is the interested users metric, and *RIU* is the really interested users metric. *VP* is the amount of transmission that users have been through. It compares the duration of the requests with the length of the full transmission, obtaining the resulting percentage. It must be taken into account that this metric is not eligible for continuous broadcasts (like conventional TV), because there are no time limitations. Although in continuous transmissions it could be applied to specific time ranges or programmes, a value of 100 should be used to calculate the *IoS*. *IU* represents the users that have been attracted by the access pages or the advertisements that have been distributed. For its calculation, the total number of different users shall be counted in the server or proxies log files. *RIU* considers all the users that, apart from being attracted by the access information, have spent certain time connected to the service. This time depends on the provider's preferences and can range from a few seconds to several hours.

5 CONCLUSION

The configuration and deployment of live-video services is an extremely complex process, due to the high resource consumption of these services, and the difficulty of transmitting continuous information over a shared data network. Nowadays, this task is mainly based on managers' experience. However, a formalization of the steps which must be followed to attain a service of quality, could improve the obtained results increasing service performance and profitability. The proposed engineering method and the expounded approach to service analysis have a direct applicability in HFC networks and they are perfectly compatible with other types of networks. It could also be the base for the development of a complete analysis and configuration methodology that could support service management tasks using production information.

ACKNOWLEDGEMENTS

This research has been financed by the network operator **Telecable** and the newspaper **La Nueva España** within the *NuevaMedia*, *TeleMedia* and *ModelMedia* projects.

REFERENCES

Chawathe, Y., 2000. *Scattercast: An Architecture for Internet Broadcast Distribution as an Infrastructure Service*, Ph.D. Dissertation, University of California at Berkley, U.S.A.

Chow, R.K.Y. and Tham, C.K., 2000. Scalable Video Delivery to Unicast Handheld-Based Clients. *Proceedings of the 2000 IEEE International Conference on Networks* (*IEEE ICON 2000*), Singapore, pp. 93-98.

Deshpande, H. et al, 2001. *Streaming Live Media over a Peer-To-Peer Network*, Technical Report, Standford University, U.S.A.

García, V.G., et al, 2003. *Redes de Acceso de Banda Ancha, Arquitectura, Prestaciones, Servicios y Evolución*, Telecable and Spanish Ministry of Science and Technology, Madrid, Spain, pp. 37-64.

Nguyen, T.P. et al, 2002. Distributed Video Streaming over the Internet. *Proceedings of Multimedia Computing and Networking (MMCN'02)*, California, U.S.A.

Ortega, A., 2000. *Variable Bit-Rate Video Coding*, in *Compressed Video over Networks*, M.-T. Sun and A. R. Reibman, Eds, Marcel Dekker, New York, U.S.A., pp. 343-382

Padmanabhan, V.N. et al, 2002. Distributing Streaming Media Content Using Cooperative Networking. *Proceedings of ACM NOSSDAV 2002*, Florida, U.S.A.

RealNetworks, 2002. *Helix Universal Server Administration Guide*, RealNetworks, Inc.

Tham, C.K. et al, 2003. Layered Coding for a Scalable Video Delivery System. *Proceedings of IEEE/EURASIP Packet Video 2003* (*PV 2003*), Nantes, France.

Turletti, T. and Bolot, J.C., 1994. Issues with Multicast Video Distribution in Heterogeneous Packet Networks. *Proceedings of 6th International Workshop on Packet Video*, Portland, U.S.A., pp. F3.1-F3.

A HARDWARE-ORIENTED ANALYSIS OF ARITHMETIC CODING – COMPARATIVE STUDY OF JPEG2000 AND H.264/AVC COMPRESSION STANDARDS

Grzegorz Pastuszak
Warsaw University of Technology, Institute of Radioelectronics, Nowowiejska 15/19, Warsaw, Poland
Email: G.Pastuszak@ire.pw.edu.pl,

Keywords: CABAC, binary arithmetic coding, H.264, MPEG-4 AVC, JPEG 2000.

Abstract: This paper provides an in-depth analysis and comparison of the arithmetic coding stages in the latest compression standards: JPEG 2000 and H.264/AVC for image and video systems, respectively. An impact of algorithm differences on hardware architecture is considered. Evaluation results show throughput requirements that real-time multimedia applications have to satisfy.

1 INTRODUCTION

Coding efficiency is the one of the most important features of all compression systems. Towards this goal, they follow a general schema of three main consecutive stages: modelling, quantization and coding. The latter stage exploits some well-known techniques, along with variable length codes and arithmetic coding, which map input symbols into binary sequences. The produced code streams achieve shorter lengths with respect to their source representation by making them dependent on occurrence probabilities, as Shannon's theorem claims. Arithmetic coders are able to attain better compression efficiency due to their property to effectively map input data onto binary sequences with fractional accuracy of lengths for entropy approximation. Adaptation to local statistics provides a path to further reduction of code stream lengths. However, these properties imply much higher computational complexity. JPEG 2000 (ISO/IEC 15444-1, 2000) and H.264/AVC (ISO/IEC 14496-10, 2003) are standards, where the Context Adaptive Binary Arithmetic Coding (CABAC) is the part of the compression scheme. Although CABAC bases on the same general principles in both standards, there are some substantial differences between them.

The comparison of the standards provided in (Marpe et al., 2003-Oct.) focuses mainly on the compression efficiency providing complexity issues rather in broad outline. CABAC algorithms, as they stand, are extensively described in related works for both JPEG2000 (Taubman et. al., 2002) and H.264/ AVC (Marpe et al., 2003-July), (Marpe et al., 2003-Sept.). In case of JPEG 2000, the bottleneck of the system arises from the entropy coding stage along with the arithmetic coder. Some optimisation methods were reported in literature and they lend themselves to the newest video compression schema. On the other hand, existing differences necessitate unique approaches. In this paper, we emphasize design details in terms of both hardware complexity and speed. Moreover, throughput evaluations are provided to find speed requirements for real-time applications.

The remainder of the paper is organized as follows: Section 2 illustrates consecutive stages of the CABAC that are deeply analysed in subsections 2.1 – 2.4. Subsection 2.5 addresses bypass mode variants. System-level conditions for arithmetic coding are given in Section 3.1. The following ones provide test conditions, evaluated requirements for processing speed and discussion combining them with hardware design methods; finally, Section 4 concludes the work.

2 MAIN STAGES OF THE CABAC

In terms of basic operations while executing the CABAC algorithm, we may distinguish four main stages. Fig. 1 shows their causal arrangement what suggest how to implement the algorithm in hardware

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 255–262.

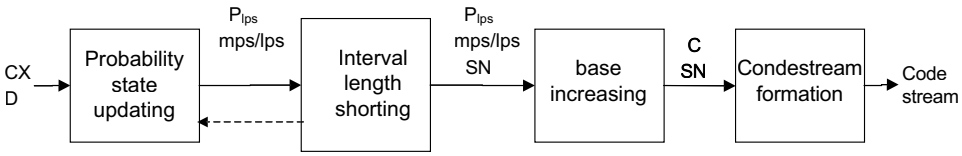


Figure 1: Division of the arithmetic coding algorithm in terms of casual relationships allows pipelined architectures. Here, Plps denotes probability estimate of LPS, SN – renormalization shift number, C – the base before renormalization, mps/lps – selects between MPS and LPS. The dashed line identifies optional feed back (JPEG2000).

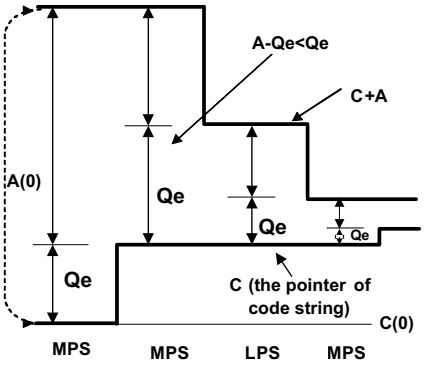


Figure 2: The interval subdivision in the JPEG2000 arithmetic coder.

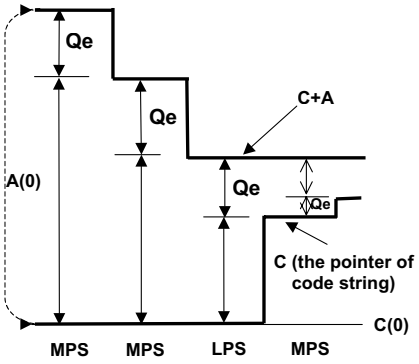


Figure 3: The interval subdivision in the H.264/AVC arithmetic coder.

to obtain a high throughput. First stage performs updating the probability state array according to pre-defined transaction rules. Each element of this state array corresponds to one of the possible input contexts, CX, and consists of two fields: the index and the most probable symbol (MPS) value. The index value identifies probability estimate of the least probable symbol (LPS). An estimate for a given context is forwarded to the second stage to subdivide the current probability interval (A) into two ones, as illustrated in Fig. 2 and Fig. 3. Depending on LPS/MPS coding, one of them is selected as a new one, and renormalized to desired range by shifting left, if needed. The third stage manages the interval base register (C - lower endpoint). This register is increased when the upper subinterval is selected as a new one. Successive renormalization shifts for the A register trigger the analogous behaviour of the C one, which releases code bits from its MSB positions. The bits are collected in the last stage into bytes and output to external functional blocks to form a final compressed stream.

2.1 Probability State Updating

In JPEG 2000 and H.264/AVC, there are respectively 19 and 399 possible contexts defined for the CABAC. Each context has an associated finite state machine conveying the index, as a 6-bit vector, which assumes 47, in JPEG 2000, and 64, in H.264/AVC, allowable values. The small number of contexts, in the image compression standard, enables hardware architectures to implement the probability state array in registers, whereas the video schema imposes using an on-chip memory to save area of an integrated circuit. In spite of this drawback, H.264/AVC exhibits more flexible properties for pipeline-oriented approaches, since probability state updating process experience no impact from the probability interval renormalization. Such dependencies exist in the case of JPEG 2000, while coding MPS. If it does not cause renormalization shifts then the probability state, pointed by the current context, remains unchanged; otherwise a

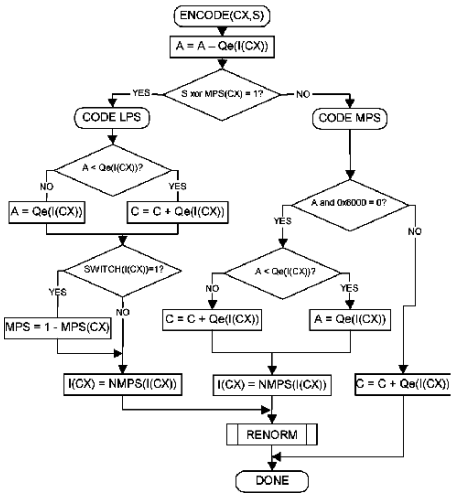


Figure 4: Arithmetic coding procedure in JPEG2000 embeds the conditional exchange of subintervals, conditional probability update for MPS.

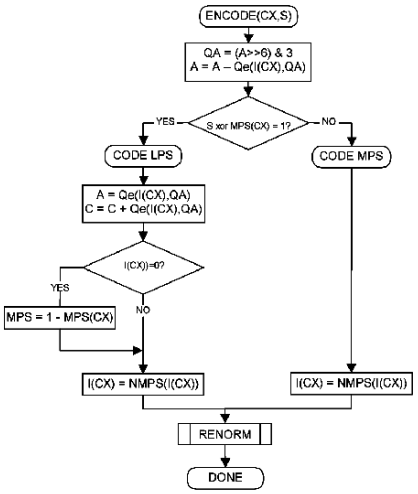


Figure 5: Arithmetic coding procedure in H.264/AVC: QA indicates two bits classifying the interval length (A) to one of four subranges to better match probability estimates. LPS is always assigned to upper subinterval, whereas MPS to the lower one.

new value is stored in accordance with the index transaction table. In the image compression schema, indices, in the probability state array, are initialised always to the same values defined in the standard specification. H.264/AVC involves quite complicated initialisation rules dependent on the quantization parameter (Qp), the frame type and a particular context number. Moreover, for INTER frames, we can select between three sets of initialisation schemas. The best choice may be done on the base of the resultant compression rate and depends on the video content. As for circuit, the initialisation rules introduce a considerable amount of both silicon and time resources. The first implication arises from the need to keep a large number of pre-defined constants in either the ROM table or combinatorial logic. The second one is caused by the necessity to check the rate for three initialisation cases, while coding INTER frames. Incorporating three CABAC engines, operating in parallel, can solve the time problem at the expense of hardware resources.

2.2 Interval Length Calculation

JPEG 2000 incorporates the 16-bits interval register, whereas H.264/AVC uses the 9-bits one. As for hardware, the increased precision improves slightly coding efficiency at the expense of resources and a longer carry chain. In considered standards, the interval length undergoes multiplication-free modifications, which realize its subdivision into two disjoint ranges. Since multiplication products are replaced by their tabled approximations, we would deal with some losses in coding efficiency (up to 3%). To mitigate this drawback, both compression schemes utilize different approaches. In JPEG 2000, the CABAC executes the conditional exchange, which ensures that the larger and smaller subintervals are always assigned to MPS and LPS, respectively. A fast hardware implementation translates this feature to a comparator (A<2Qe?) driving, together with MPS/LPS signal, the selection of an appropriate subinterval. The associated latency of this operation matches that in subtraction carry chain (A-Qe), thereby, the additional circuit should not deteriorate working frequency. H.264/AVC employs another solution to improve coding efficiency. Prior to subdivision (by subtraction), the interval is classified to one of four ranges based on its two bits located just after the MSB bit (which should equa1 1). Each range corresponds to a separate set of probability estimates to better approximate a multiplication product (A*Qe). Hence, given a set, the index points an appropriate estimate ready for either the subtraction or conditional interval reloading. A circuit, implementing the H.264/AVC arithmetic coder, has to provide all

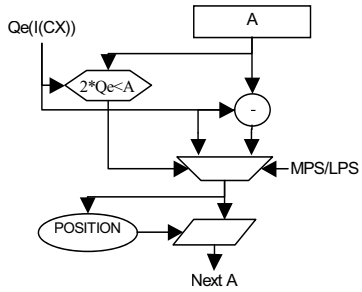


Figure 6: The A-interval subdivision circuit in the JPEG2000 arithmetic coder.

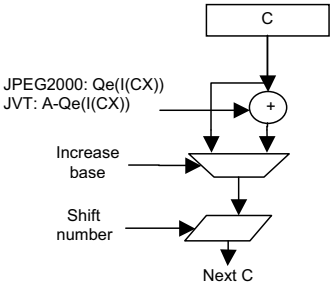


Figure 8: The C-base updating circuit able to process one symbol per clock cycle.

necessary probability estimate tables what amounts to a relevant combinatorial logic. The generated estimates should be available before beginning subinterval calculation to minimize propagation delays. Therefore, the pipelined architecture should determine these estimates in a preceding stage. However, selection between four ones, depended on the interval bits, must be held at the second stage what affects negatively timing. Determining estimates in advance fails in case of JPEG 2000 because of conditional state updating, as described above. On the other hand, there is no need for the selection. Moreover, less significant bits of estimates vary weakly over all indexes, what drives the logic synthesizer to simplify circuits, thus, shorting critical paths. After all, the interval value is renormalized by means of the shifting circuit. Fig. 6 and 7 depict interval length calculation circuits able to process one symbol per clock cycle for considered standards.

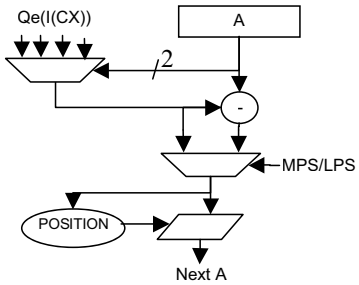


Figure 7: The A-interval subdivision circuit in the H.264/AVC arithmetic coder.

2.3 Interval Base Calculation

The base register in H.264/AVC needs 10 bits to be implemented. The location of subintervals is inverted with respect to JPEG 2000, i.e., in video compression scheme the upper and the lower correspond to LPS and MPS coding, respectively. Most notably, while subtracting the probability estimate from the interval length (LPS coding), the base has to be increased by the subtraction product. This dependency may affect negatively clock rate. Locating the base circuit in a separate pipeline stage gives somewhat shorter critical paths. As a consequence of similar bit counts in carry chains (9 and 10), the delays of pipelined circuits match one another. This observation holds for renormalization units due to an identical shift number submitted to both. The JPEG 2000 arithmetic coder keeps the lower bound of the interval in the 28 bits register. It experiences the analogous operations as the H.264/AVC counterpart. From hardware perspective, the main difference arises from their sizes what reflects on the latency in carry chains. In order to shorten critical paths in JPEG 2000, it is viable to divide the base register into two parts and place them in consecutive pipeline stages. Such rearrangement refers also to relevant combinatorial logic. Another distinguishing feature of the image compression coder lays in the meaning of the oldest part of the lower bound register, which supports carry propagation and code data releasing. The H.264/AVC version applies a different solution to these problems, so that, it removes the need to extend the base register towards MSB bits. The typical circuit managing the base is shown in Fig. 8.

2.4 Code Stream Generation

In the H.264/AVC arithmetic coder, output bits are released from the second MSB position of the base register after each single renormalization shift. In order to solve the problem of carry propagation, ones, encountered in series, are counted without outputting. Occurrence of the carry, indicated by the MSB bit of the base register, activates releasing binary one followed by a number of zeros. Otherwise, the inverted version of such sequence appears as an outcome after encountering a zero bit. This procedure requires the use of a counter signalling a total of outstanding bits. Its precision should match a maximal possible code stream length to prevent overflow when dealing with an extremely long series. The produced bits are assembled into bytes and released. It may happen that more then one byte has to be output due to a large number of outstanding bits. Provided the CABAC accepts one symbol per clock cycle, the design must adjust this rate to irregular code byte generation conditions by inserting wait-states. The JPEG 2000 arithmetic coder is free to that problem since at most two code bytes can appear as an outcome after processing one symbol. The algorithm imposes the need to keep the last generated byte in the buffer ready to complete carry. If there is the 0xFF byte, the control logic inserts one stuffing bit into the MSB position of the following byte. This bit assumes the zero value to trap a carry. A dedicated down-counter points to bits in the base register that have not been released so far. In terms of higher performances, both compression standards find a separate pipeline stage to make the code stream generation adequate.

2.5 Bypass Mode

The CABAC in H.264/AVC provides the bypass mode, which, against the regular one, assumes uniform probability distribution of submitted symbols. Hence, it skips the probability state updating routine. Since related symbols contribute to the same code stream as in regular mode, it is natural to use the same resources with their timing constraints. The interval register remains unchanged in bypass mode. This property, in conjunction with skipping the probability adaptation, gives an opportunity to process bypassed symbols and regular ones in parallel. The probability estimates are obtained by single shifting right (division by 2) the interval value. Therefore, we must append one bit to represent estimate accurately. In case of JPEG 2000, the bypass mode forwards symbols directly to the output stream without arithmetic encoding. As for hardware, this approach allows increasing the throughput to a rate determined by the bit-plane coder performances, which submits input data to the CABAC module. However, the total improvement is not so significant since the standard enables the bypass mode for some coding passes over lowest bit planes.

3 EVALUATION

3.1 System-Level Constraints

Since the CABAC in H.264/AVC produces the single code stream for an entire slice, all necessary context-symbol data have to be applied to the one functional block. Thus, its speed determines the overall performances of the coder when input data are received continuously. Lower bit-rates decrease demands for throughput. Using rate-distortion optimisation for each macro-block improves quality at the same compression ratio. To obtain rates, we need to carry out arithmetic coding (when used) for all possible coding modes. As a consequence, it burdens the CABAC with a large number of computations and may lead to timing constraints for the encoder. JPEG 2000 supports entropy-coding parallelism by independent analysing rectangular blocks of coefficients in the wavelet domain. Each such code block generates a separate output stream, which can be truncated in some points to increase the compression ratio at the expense of quality losses of the reconstructed image. Moreover, a special mode drives the arithmetic coder to terminate the stream on these points. For the sake of the rate control policy, it is desired to produce more outcomes to discard their less significant parts with reference to the optimisation criteria. Thus, we need faster CABAC engines to support this property.

3.2 Evaluation Conditions

Evaluations have been conducted for some video sequences taking into account the number of binary symbols submitted to the arithmetic coder in both standards. We examined test cases relating to CIF and QCIF resolutions. As reference software for image compression schema, we have employed JJ2000 version 5.1 adapted to support video material as Motion JPEG 2000 (ISO/IEC 15444-3, 2002). To get characteristics following options have been used: no tiling, five levels of wavelet decompositions, 9/7 wavelet filter, code block size of 64 x 64 samples, regular coding mode, single quality layer. Explicit quantization by step size has enabled to vary both

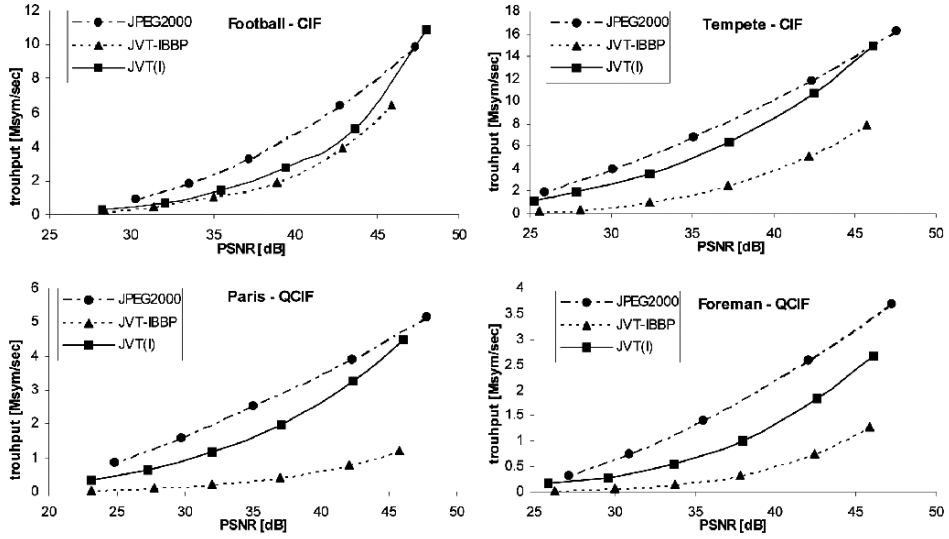


Figure 9. Averaged throughput requirements for H.264/AVC (JVT) and Motion JPEG2000.

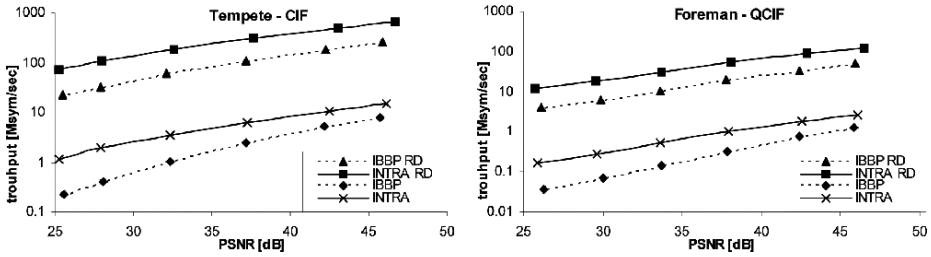


Figure 10. Averaged throughput requirements for H.264/AVC (JVT) using RD optimization or not.

quality and rate. Additionally, the special case, when no quantization is present, has been checked to demonstrate the most critical conditions, when total losses in quality of the reconstructed images results only from the rate-control policy. For encoding video material, the Joint Model (JM) of the Joint Video Team (JVT), software version 7.4 has been used. We have explored constraints when taking advantage of the RD-optimisation (rate-distortion) or not. All tests have skipped the rate controller to determine quality losses by explicit quantization parameter, Qp. The list of other settings has been as follows: only one slice per picture, 2 reference frames, full search of motion vectors, in IBBP mode: I-frames every 15th, 2 B-frames between I and/or P. All evaluations assumed frame rate of 30 Hz.

3.3 Evaluation Results

Fig. 9 shows throughput (number of coded symbols) versus quality expressed as average PSNR (distortion) of the luminance component over all frames in a given sequence. The average PSNR of the chrominance components (U, V) has been adjusted with reference to that of the luminance one to obtain approximately the same differences for both compression schemas. This objective has been achieved by first producing curves for JPEG 2000, and then, iteratively varying the quantization parameter offset for chrominance components in H.264/AVC reference model. As expected, the throughput requirements increase significantly if we

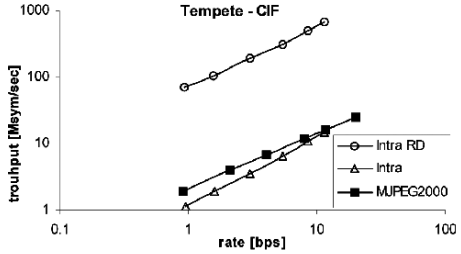


Figure 11. Averaged throughput requirements for H.264/AVC (JVT) and Motion JPEG2000 versus bit-rate.

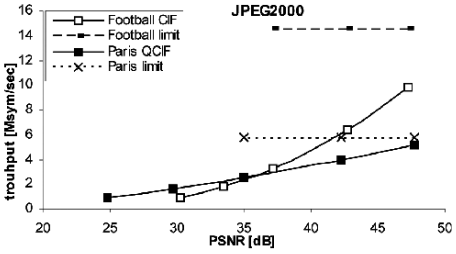


Figure 12. Averaged throughput requirements for Motion JPEG2000 with indicated maximal limits when no quantization.

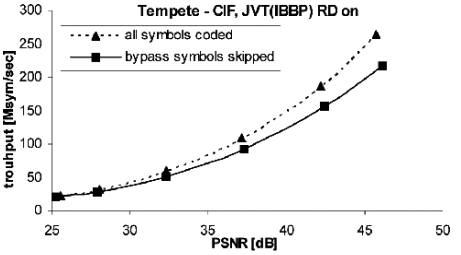


Figure 13. Averaged throughput requirements for H.264/AVC diminish when symbols obeying uniform distribution are skipped.

want to decrease distortions. Without RD optimisation, the video compression schema takes less time to accomplish the arithmetic coding routine than JPEG 2000. The INTRA mode of H.264/AVC demands a little smaller throughput then JPEG 2000 for the quality range of interest to most video applications. However, at higher qualities, both standards need similar processing speed of the CABAC block with the tendency to favour the JPEG 2000. Of course, the use of INTER mode makes H.264/AVC the best solution in terms of both compression ratio and throughput of the entropy coder. On the other hand, the compressed stream must embed the INTRA frames, so, their temporal impact on the latency of the arithmetic coder should be taken into account. Using RD optimisation in H.264/AVC increases demands for the CABAC module by about two orders of magnitude (see Fig. 10). This computation growth is necessary only to select the best mode for macroblock in the sense of Lagrange's minimization. The opportunity for parallelism arises, since a macroblock can be processed for different modes at the same time. As depicted in Fig. 11, the computation demands for the CABAC engine are almost proportional to a given rate in both standards. So, on the base of imposed bandwidth or storage limitations, we can estimate desired throughput of the arithmetic coder.

3.4 Discussion

With the above observations, we re-examine the pipelined structure of the arithmetic coders for real time performances. Since the context generator in JPEG 2000 occupies a relatively large silicon area of the integrated circuits (due to coefficient memory), it is payable to optimize the throughput of the CABAC unit rather then to duplicate entropy coding paths including both modules. Such single path, embedding pipeline architecture able to process one symbol per clock cycle, yields speed to target from CIF sequences (4:2:0 – 4:4:4, 30 frames per second) provided regular and lossless mode (Hsiao et. al, 2002), (Lian et al, 2003), (Li et al, 2002), (Fang et al, 2002). Approaches based on sequential arrangement attain worse results in spite of higher clock rates (Andra K et al, 2003). The lossy compression allows higher throughputs, as shown in Fig. 12. Of course, the exact performances depend primarily on the technology of an integrated circuits and efforts spent to minimize critical paths. To speed up the entropy coding, we can use parallel processing paths and/or modify the architecture to process two or more symbols per clock cycle.

In H.264/AVC encoder without RD optimisations, the single CABAC engine, complying with pipeline arrangement (like in JPEG 2000) able to process one symbol per clock cycle, can easy support PAL and NTSC standards in the compression range of most interest. Moreover, it makes possible to target HDTV at lower bit-rates (low quality). However, we must remember that the throughput of the whole video coder depends on the motion estimation unit rather then the entropy

coding stage. Taking advantage of the RD optimisation finds the arithmetic coder to become another bottleneck of the system. As mentioned above, employing several engines in parallel mitigates timing constraints at the cost of silicon resources. This can be realized in two ways. The first assumes dividing a frame into a number of slices and assigning one CABAC unit to each of them. Prior to checking all coding scenarios, we have to save the states of internal registers, and then, encode a macroblock in series for various modes staring from the same state. The second way determines rates simultaneously in separate coding units. One can combine these approaches as well. Since some symbols obey uniform distribution, they induce extension of the output stream by one bit. So, when we want to estimate rates, it is enough to count them without submitting to the arithmetic coder. Fig. 13 depicts the difference in throughput of the CABAC, while benefiting from this opportunity.

4 CONCLUSIONS

The analysed arithmetic coding algorithms proves to comply with the general schema of the pipeline architecture design. Corresponding stages exhibit some variants of the CABAC concepts requiring different approaches to minimize critical paths. The H.264/AVC version can achieve higher working frequencies than the JPEG 2000 one due to smaller sizes of the key registers. Owing to the latter supports entropy coding parallelism, it can achieve high performance, but a hardware designer should primarily sophisticate the single entropy channel to save a total of silicon area. Special attention must be paid to optimise the CABAC unit in H.264/AVC, when RD optimisation is on, including parallel encoding engines, counting bypassed symbols, and minimizing critical paths. Without RD enhancements, the throughput of the single CABAC gives opportunity even to target HDTV.

ACKNOWLEDGEMENTS

The work presented was developed within activities of VISNET, the European Network of Excellence, (http://www.visnet-noe.org), founded under the European Commission IST 6FP programme.

REFERENCES

ISO/IEC 15444-1, 2000, *JPEG 2000 Part I Final Committee Draft Version 1.0*.

ISO/IEC 15444-3, 2002, *Motion – JPEG 2000 Part 3*.

ISO/IEC 14496-10*,* 2003, *ITU-T Recommendation H.264 and ISO/IEC 14496-10 MPEG-4 Part 10, Advanced Video Coding (AVC).*

Taubman D. S. and Marcellin M. W., 2002, *JPEG2000: Image Compression Fundamentals, Standard and Practice*. Norwell, MA: Kluwer.

Andra K., Chakrabarti C., and Acharya T., 2003, "A high performance JPEG2000 architecture," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 3, pp. 209–218.

Hsiao Y-T, Lin H-D, Lee K-B and Jen C-W*,* 2002, *"*High Speed Memory Saving Architecture for the Embedded Block Coding in JPEG 2000".

Lian C.-J., Chen K.-F., Chen H.-H., and Chen L.-G., 2003, "Analysis and architecture design of block-coding engine for EBCOT in JPEG 2000," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 3, pp. 219–230.

Li Y., Aly R.E., Wilso B. and Bayoumi *M.A.*, 2002, "Analysis and Enhancement for EBCOT in high speed JPEG 2000 Architectures," The 45th Midwest Symposium on Circuits and Systems, 2002. MWSCAS-2002.

Fang H-C, Wang T-C, Lian C-J, Chang T-H and Chen L-G*,* 2002, *"*High Speed Memory Efficient EBCOT Architecture for JPEG2000".

Marpe D., Schwarz H, and Wiegand T., 2003-July, "Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard", *IEEE Transactions on Circuits and Systems for Video Technology*.

Marpe D. and Wiegand T., 2003-Sept., "A highly Efficient Multiplication-Free Binary Arithmetic Coder and Its Application in Video Coding", *Proc. IEEE International Conference on Image Processing (ICIP 2003)*, Barcelona, Spain.

Marpe D., George V., Cycon H.L., Barthel K.U., 2003-Oct, "Performance evaluation of Motion-JPEG2000 in comparison with H.264/AVC operated in pure intra coding mode", *Proc. SPIE on Wavelet Applications in Industrial Processing*, Rhode Island, USA.

AUDIO WATERMARKING QUALITY EVALUATION

Andrés Garay Acevedo
Georgetown University, Washington, DC, USA
Email: ag66@georgetown.edu

Keywords: Audio watermarking, benchmarking, perceptual model.

Abstract: The recent explosion of the Internet as a collaborative medium has opened the door for people who want to share their work. Nonetheless, the advantages of such an open medium can pose very serious problems for authors who do not want their works to be distributed without their consent. As new methods for copyright protection are devised, expectations around them are formed and sometimes improvable claims are made. This paper covers one such technology: audio watermarking. First, the framework is set for the objective measurement of such techniques. After this, the remainder of the document proposes a test and a set of metrics for thorough benchmarking of audio watermarking schemes. The development of such a benchmark constitutes a first step towards the standardization of the requirements and properties that such systems should display.

1 INTRODUCTION

A watermarking process can be modeled as a communication process. In fact, this assumption is used throughout this paper, as it will prove to be beneficial at a later stage. A more detailed description of this model can be found in (Cox, Miller, & Bloom, 2002).

In this framework, watermarking is viewed as a transmission channel through which the watermark message is communicated. Here the cover work is just part of the channel. This is depicted in figure 1, adapted from (Cox et al., 2002).

The embedding process consists of two steps. First, the watermark message m is mapped into an added pattern[1] W_a, of the same type and dimension as the cover work A. When watermarking audio, the watermark encoder produces an audio signal. This mapping may be done with a watermark key K. Next, W_a is embedded into the cover work in order to produce the watermarked audio file A'.

After the pattern is embedded, the audio file is processed in some way. This is modeled as the addition of noise to the signal, which yields a noisy work A'_n. The types of processing performed on the work will be discussed later, as they are of no importance at this moment. However, it is important to state the presence of noise, as any transmission medium will certainly induce it.

The watermark detector performs a process that is dependant on the type of watermarking scheme. If the decoder is a *blind or public decoder* , then the original audio file A is not needed during the recovery process, and only the key K is used in order to decode a watermark message m_n.

[1] This pattern is also known as a pseudo-noise (PN) sequence. Even though the watermark message and the PN-sequence are different, it is the later one we refer to as the watermark W

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 263–274.

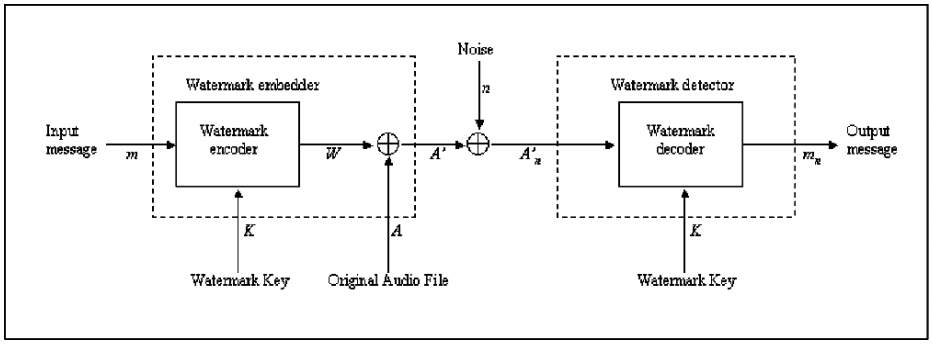


Figure 1: Watermark communication process.

Another possibility is for the detector to be *informed* . In this case, the original audio cover A must be extracted from A'_n in order to yield W_n, prior to running the decoding process. In addition, a confidence measure can be the output of the system, rather than the watermark message.

In order to measure the quality of a watermarking scheme, one can perform a different test at several points of the communication process. In fact, this is exactly what is proposed on this document. These points are namely the sending and receiving ends, and the communication channel. Moreover, at these points specific actors (with different concerns about the technology) take part in the process. The rest of this document addresses these concerns, as it outlines three specific subtests for evaluating watermarking systems. Finally, these tests are combined in order to produce a final watermarking test score.

2 MEASURING FIDELITY

Artists, and digital content owners in general, have many reasons for embedding watermarks in their copyrighted works. These reasons have been stated on various occasions. However, there is a big risk in performing such an operation, as the quality of the musical content might be degraded to a point where its value is diminished. Fortunately, the opposite is also possible and, if done right, digital watermarks can add value to content (Acken, 1998).

Content owners are generally concerned with the degradation of the cover signal quality, even more than users of the content (Craver, Yeo, & Yeung, 1998). They have access to the unwatermarked content with which to compare their audio files. Moreover, they have to decide between the amount of tolerance in quality degradation from the watermarking process and the level of protection that is achieved by embedding a stronger signal. As a restriction, an embedded watermark has to be detectable in order to be valuable.

Given this situation, it becomes necessary to measure the impact that a marking scheme has on an audio signal. This is done by measuring the *fidelity* of the watermarked audio signal *A.'*

As fidelity refers to the similitude between an original and a watermarked signal, a statistical metric must be used. Such a metric will fall in one of two categories: difference metrics or correlation metrics.

Difference metrics, as the name states, measure the difference between the undistorted original audio signal A and the distorted watermarked signal A'. In the case of digital audio, the most common difference metric used for quality evaluation of watermarks is the signal to noise ratio (SNR). This is usually measured in decibels (dB), so SNR(dB) = 10 $\log_{10}$ (SNR).

The signal to noise ratio, measured in decibels, is defined by the formula

$$SNR(dB) = 10\log_{10}\frac{\sum_n A_n^2}{\sum_n (A_n - A'_n)^2} \quad , \qquad (1)$$

where A_n corresponds to the n^{th} sample of the original audio file A, and A'_n to the n^{th} sample of the watermarked signal A'. This is a measure of quality that reflects the quantity of distortion that a watermark imposes on a signal (Gordy & Burton, 2000).

Another common difference metric is the peak signal to noise ratio (PSNR), which measures the maximum signal to noise ratio found on an audio signal. A description of the PSNR, along with some other difference metrics found on the literature is presented on (Kutter & Hartung, 2000; Kutter & Petitcolas, 1999).

Although the tolerable amount of noise depends on both the watermarking application and the characteristics of the unwatermarked audio signal, one could expect to have perceptible noise distortion for SNR values of 35dB (Petitcolas & Anderson, 1999).

Correlation metrics measure distortion based on the statistical correlation between the original and modified signals. They are not as popular as the difference distortion metrics, but it is important to state their existence.

For the purpose of audio watermark benchmarking, the use of the signal to noise ratio should be used to measure the fidelity of the watermarked signal with respect to the original. This decision follows most of the literature that deals with the topic (Gordy & Burton, 2000; Kutter & Petitcolas, 1999, 2000; Petitcolas & Anderson, 1999). Nonetheless, in this measure the term noise refers to statistical noise, or a deviation from the original signal, rather than to perceived noise on the side of the hearer. This result is due to the fact that the SNR is not well correlated with the human auditory system (Kutter & Hartung, 2000). Given this characteristic, the effect of perceptual noise needs to be addressed later.

In addition, when a metric that outputs results in decibels is used, comparisons are difficult to make, as the scale is not linear but rather logarithmic. This means that it is more useful to present the results using a normalized quality rating. The ITU-R Rec. 500 quality rating is perfectly suited for this task, as it gives a quality rating on a scale of 1 to 5 (Arnold, 2000; Piron et al., 1999). Table 1 shows the rating scale, along with the quality level being represented.

Table 1: ITU-R Rec. 500 quality rating.

Rating	Impairment	Quality
5	Imperceptible	Excellent
4	Perceptible, not annoying	Good
3	Slightly annoying	Fair
2	Annoying	Poor
1	Very annoying	Bad

The fidelity of the watermarked signal is computed by using the formula

$$Fidelity = \frac{5}{1 + N * SNR}, \quad (2)$$

where N is a normalization constant and SNR is the measured signal to noise ratio.

2.1 Data Payload

The fidelity of a watermarked signal depends on the amount of embedded information, the strength of the mark, and the characteristics of the host signal. This means that a comparison between different algorithms must be made under equal conditions. That is, while keeping the payload fixed, the fidelity must be measured on the same audio cover signal for all watermarking techniques being evaluated.

However, the process just described constitutes a single measure event and will not be representative of the characteristics of the algorithms being evaluated, as results can be biased depending on the chosen parameters. For this reason, it is important to perform the tests using a variety of audio signals, with changing size and nature (Kutter & Petitcolas, 2000). Moreover, the test should also be repeated using different keys.

The amount of information that should be embedded is not easy to determine, and depends on the application of the watermarking scheme. In (Kutter & Petitcolas, 2000) a message length of 100 bits is used on their test of image watermarking systems as a representative value. However, some secure watermarking protocols might need a bigger payload value, as the watermark W could include a cryptographic signature for both the audio file A, and the watermark message m in order to be more secure (Katzenbeisser & Veith, 2002). Given this, it is recommended to use a longer watermark bitstream for the test, so that a real world scenario is represented. A watermark size of 128 bits is big enough to include two 56-bit signatures and a unique identification number that identifies the owner.

3 MEASURING ROBUSTNESS

Watermarks have to be able to withstand a series of signal operations that are performed either intentionally or unintentionally on the cover signal and that can affect the recovery process. Given this, watermark designers try to guarantee a minimum level of *robustness* against such operations. Nonetheless, the concept of robustness

is ambiguous most of the time and thus claims about a watermarking scheme being robust are difficult to prove due to the lack of testing standards (Craver, Perrig, & Petitcolas, 2000).

By defining a standard metric for watermark robustness, one can then assure fairness when comparing different technologies. It becomes necessary to create a detailed and thorough test for measuring the ability that a watermark has to withstand a set of clearly defined signal operations. In this section these signal operations are presented, and a practical measure for robustness is proposed.

3.1 How to Measure

Before defining a metric, it must be stated that one does not need to erase a watermark in order to render it useless. It is said that a watermarking scheme is robust when it is able to withstand a series of attacks that try to degrade the quality of the embedded watermark, up to the point where it's removed, or its recovery process is unsuccessful. This means that just by interfering with the detection process a person can create a successful attack over the system, even unintentionally.

However, in some cases one can overcome this characteristic by using error-correcting codes or a stronger detector (Cox et al., 2002). If an error correction code is applied to the watermark message, then it is unnecessary to entirely recover the watermark W in order to successfully retrieve the embedded message m. The use of stronger detectors can also be very helpful in these situations.

Given these two facts, it makes sense to use a metric that allows for different levels of robustness, instead of one that only allows for two different states (the watermark is either robust or not). With this characteristic in mind, the basic procedure for measuring robustness is a three-step process, defined as follows:

For each audio file in a determined test set embed a random watermark W on the audio signal A, with the maximum strength possible that doesn't diminish the *fidelity* of the cover below a specified minima (Petitcolas & Anderson, 1999).

Apply a set of relevant signal processing operations to the watermarked audio signal A'.

Finally, for each audio cover, extract the watermark W using the corresponding detector and measure the success of the recovery process.

Some of the early literature considered the recovery process successful only if the whole watermark message m was recovered (Petitcolas, 2000; Petitcolas & Anderson, 1999). This was in fact a binary robustness metric. However, the use of the *bit-error rate* has become common recently (Gordy & Burton, 2000; Kutter & Hartung, 2000; Kutter & Petitcolas, 2000), as it allows for a more detailed scale of values. The *bit-error rate* (BER) is defined as the ratio of incorrect extracted bits to the total number of embedded bits and can be expressed using the formula

$$BER = \frac{100}{l}\sum_{n=0}^{l-1}\begin{cases}1, & W'_n = W_n \\ 0, & W'_n \neq W_n\end{cases}, \qquad (3)$$

where l is the watermark length, W_n corresponds to the n^{th} bit of the embedded watermark and W'_n corresponds to the n^{th} bit of the recovered watermark. In other words, this measure of robustness is the certainty of detection of the embedded mark (Arnold, 2000).

The three-step procedure just described should be repeated several times, since the embedded watermark W is randomly generated and the recovery can be successful by chance (Petitcolas, 2000).

Up to this point no details have been given about the signal operations that should be performed in the second step of the robustness test. These are now presented.

3.2 Audio Restoration Attack

In audio restoration the recording is digitized and then analyzed for degradations. After these degradations have been localized, the corresponding samples are eliminated. Finally, the missing samples are recreated by interpolating the signal using the remaining samples.

One can assume that the audio signal is the product of a stationary autoregressive (AR) process of finite order (Petitcolas & Anderson, 1998). With this assumption in mind, one can use an audio segment to estimate a set of AR parameters and then calculate an approximate value for the missing samples. Both of the estimates are calculated using a least-square minimization technique.

Using the audio restoration method just described one can try to render a watermark undetectable by processing the marked audio signal A'. The process is as follows: First divide the audio signal A' into N blocks of size m samples each. A value of m=1000 samples has been proposed in the literature (Petitcolas & Anderson,

1999). A block of length l is removed from the middle of each block and then restored using the AR audio restoration algorithm. This generates a reconstructed block also of size m. After the N blocks have been processed they are concatenated again, and an audio signal B' is produced. It is expected that B' will be closer to A than to A' and thus the watermark detector will not find any mark in it.

3.3 Invertibility Attack

When resolving ownership cases in court, the disputing parties can both claim that they have inserted a valid watermark on the audio file, as it is sometimes possible to embed multiple marks on a single cover signal. Clearly, one mark must have been embedded before the other.

The ownership is resolved when the parties are asked to show the original work to court. If Alice has the original audio file A, which has been kept stored in a safe place, and Mallory has a counterfeit original file $\tilde{A}$ which has been derived from A, then Alice can search for her watermark W in Mallory's file and will most likely find it. The converse will not happen, and the case will be resolved (Craver et al., 2000). However, an attack to this procedure can be created, and is known as an *invertibility attack*.

Normally the content owner adds a watermark W to the audio file A, creating a watermarked audio file $A' = A+W$, where the sign "+" denotes the embedding operation. This file is released to the public, while the original A and the watermark W are stored in a safe place. When a suspicious audio file $\tilde{A}$ appears, the difference $\tilde{W} = \tilde{A} - A$ is computed. This difference should be equal to W if A' and $\tilde{A}$ are equal, and very close to W if $\tilde{A}$ was derived from A'. In general, a correlation function $f(W, \tilde{W})$ is used to determine the similarity between the watermark W and the extracted data $\tilde{W}$. This function will yield a value close to 1, if W and $\tilde{W}$ are similar.

However, Mallory can do the following: she can subtract (rather than add) a second watermark $\hat{w}$ from Alice's watermarked file A', using the inverse of the embedding algorithm. This yields an audio file $\hat{A} = A' - \hat{w} = A + W - \hat{w}$, which Mallory can now claim to be the original audio file, along with $\hat{w}$ as the original watermark (Craver, Memon, Yeo, & Yeung, 1998).

When the two originals are compared in court, Alice will find that her watermark is present in Mallory's audio file, since $\hat{A} - A = W - \hat{w}$ is calculated, and $f(W - \hat{w}, W) \approx 1$. However, Mallory can show that when $A - \hat{A} = \hat{w} - W$ is calculated, then $f(\hat{w} - W, \hat{w}) \approx 1$ as well. In other words, Mallory can show that her mark is also present in Alice's work, even though Alice has kept it locked at all times (Craver, Memon, & Yeung, 1996; Craver, Yeo et al., 1998). A deadlock is thus created (Craver, Yeo et al., 1998; Pereira, Voloshynovskiy, Madueño, Marchand-Maillet, & Pun, 2001).

This attack is a clear example of how one can render a mark unusable without having to remove it, by exploiting the invertibility of the watermarking method. Such an attack can be prevented by using a non-invertible cryptographic signature in the watermark W; that is, using a secure watermarking protocol (Katzenbeisser & Veith, 2002; Voloshynovskiy, Pereira, Pun, Eggers, & Su, 2001).

3.4 Specific Attack on Echo Watermarking

The echo watermarking technique (Johnson & Katzenbeisser, 2000) can be easily "attacked" simply by detecting the echo and then removing the delayed signal by inverting the convolution formula that was used to embed it. However, the problem consists of detecting the echo without knowing the original signal and the possible delay values. This problem is refered to as *blind echo cancellation*, and is known to be difficult to solve (Petitcolas, Anderson, & G., 1998). Nonetheless, a practical solution to this problem appears to lie in the same function that is used for echo watermarking extraction: *cepstrum autocorrelation*. Cepstrum analysis, along with a brute force search can be used together to find the echo signal in the watermarked audio file A'.

A detailed description of the attack is given by Craver et al. (Craver et al., 2000), and the idea is as follows: If we take the power spectrum of $A'(t) = A(t) + \alpha A(t - \Delta t)$, denoted by Φ and then calculate the logarithm of Φ, the amplitude of the delayed signal can be augmented using an autocovariance function over the power spectrum $\Phi'(\ln(\Phi))$. Once the amplitude has been increased, then the "hump" of the signal becomes more visible and the value of the delay Δt can be determined (Petitcolas et al., 1998).

3.5 Collusion Attack

A collusion attack, also known as *averaging,* is especially effective against basic fingerprinting schemes. The basic idea is to take a large number of watermarked copies of the same audio file, and

average them in order to produce an audio signal without a detectable mark (Craver et al., 2000; Kirovski & Malvar, 2001).

Another possible scenario is to have copies of multiple works that have been embedded with the same watermark. By averaging the sample values of the audio signals, one could estimate the value of the embedded mark, and then try to subtract it from any of the watermarked works. It has been shown that a small number (around 10) of different copies are needed in order to perform a successful collusion attack (Voloshynovskiy, Pereira, Pun et al., 2001). An obvious countermeasure to this attack is to embed more than one mark on each audio cover, and to make the marks dependant on the characteristics of the audio file itself (Craver et al., 2000).

3.6 Signal Diminishment Attacks and Common Processing Operations

Watermarks must be able to survive a series of signal processing operations that are commonly performed on the audio cover work, either intentionally or unintentionally. Any manipulation of an audio signal can result in a successful removal of the embedded mark. Furthermore, the availability of advanced audio editing tools on the Internet, such as Audacity (Dannenberg & Mazzoni, 2002), implies that these operations can be performed without an extensive knowledge of digital signal processing techniques. The removal of a watermark by performing one of these operations is known as a signal diminishment attack, and probably constitutes the most common attack performed on digital watermarks (Meerwald & Pereira, 2002).

Given this, a set of the most common signal operations must be specified, and watermark resistance to these must be evaluated. Even though an audio file will most likely not be subject to all the possible operations, a thorough list is necessary. Defining which subset of these operations is relevant for a particular watermarking scheme is a task that needs to be done; however, this will be addressed later.

The signal processing operations are classified into different groups, according to the presentation made in (Petitcolas et al., 2001). These are:

Dynamics. These operations change the loudness profile of the audio signal.

Filter. Filters cut off or increase a selected part of the audio spectrum.

Ambience. These operations try to simulate the effect of listening to an audio signal on a room.

Conversion. Digital audio files are nowadays subject to format changes. These changes might induce significant quantization noise, as no conversion is perfect.

Lossy compression algorithms are becoming popular, as they reduce the amount of data needed to represent an audio signal. This can pose a serious problem to some watermarking schemes, as they sometimes will hide the watermark exactly in imperceptible regions.

Noise can be added in order to remove a watermark. This noise can even be imperceptible, if it is shaped to match the properties of the cover signal.

Modulation effects like vibrato, chorus, amplitude modulation and flanging are not common post-production operations. However, they are included in most of the audio editing software packages and thus can be easily used in order to remove a watermark.

Time stretch and *pitch shift*. These operations either change the length of an audio passage without changing its pitch, or change the pitch without changing its length in time.

Sample permutations. This group consists of specialized algorithms for audio manipulation, such as the attack on echo hiding just presented. Dropping of some samples in order to misalign the watermark decoder is also a common attack to spread-spectrum watermarking techniques.

It is not always clear how much processing a watermark should be able to withstand. That is, the specific parameters of the diverse filtering operations that can be performed on the cover signal are not easy to determine. In general terms one could expect a marking scheme to be able to survive several processing operations up to the point where they introduce annoying audible effects on the audio work. However, this rule of thumb is still too vague.

Fortunately, guidelines and minimum requirements for audio watermarking schemes have been proposed by different organizations such as the Secure Digital Music Initiative (SDMI), International Federation of the Phonographic Industry (IFPI), and the Japanese Society for Rights of Authors, Composers and Publishers (JASRAC). These guidelines constitute the baseline for any robustness test. In other words, they describe the minimum processing that an audio watermark should be able to resist, regardless of their intended application.

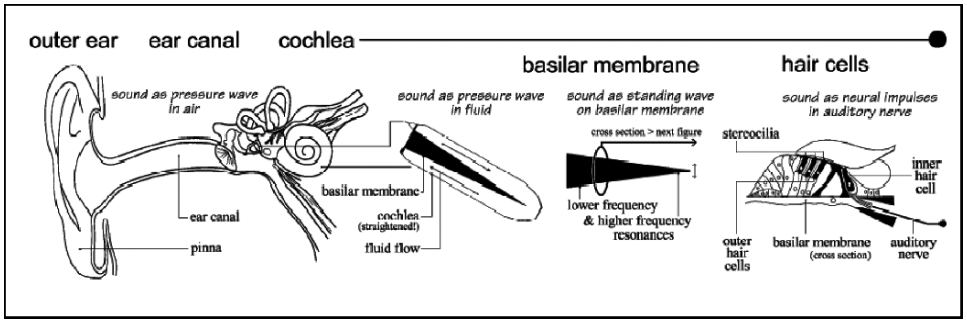


Figure 2: Overview of the Human Auditory System (HAS).

4 MEASURING PERCEPTIBILITY

Digital content consumers are aware of many aspects of emerging watermarking technologies. However, only one prevails over all of them: users are concerned with the appearance of perceptible (audible) artifacts due to the use of a watermarking scheme. Watermarks are supposed to be imperceptible (Cox et al., 2002). Given this fact, one must carefully measure the amount of distortion that the listener will perceive on a watermarked audio file, as compared to its unmarked counterpart; that is, the *perceptibility* of the watermark. Formal listening tests have been considered the only relevant method for judging audio quality, as traditional objective measures such as the signal-to-noise ratio (SNR) or total-harmonic-distortion (THD) have never been shown to reliably relate to the perceived audio quality, as they can not be used to distinguish inaudible artifacts from audible noise (ITU, 2001; Kutter & Hartung, 2000; Thiede & Kabot, 1996). There is a need to adopt an objective measurement test for perceptibility of audio watermarking schemes.

4.1 The Human Auditory System (HAS)

Figure 2, taken from (Robinson & Hawksford, 1999), presents the physiology of the human auditory system. Each one of its components is now described.

The *pinna* directionally filters incoming sounds, producing a spectral coloration, known as Head Related Transfer function (or HRTF). This function enables human listeners to localize the sound source in three dimensions. The *ear canal* filters the sound, attenuating both low and high frequencies. As a result, a resonance arises around 5 kHz. After this, small bones known as the *timpanic membrane* (or ear drum), *malleus* and *incus* transmit the sound pressure wave through the middle ear. The outer and middle ear perform a band pass filter operation on the input signal.

The sound wave arrives at the fluid-filled *cochlea*, a coil within the ear that is partially protected by a bone. Inside the cochlea resides the *basilar membrane* (BM), which semi-divides it. The basilar membrane acts as a spectrum analyzer, as it divides the signal into frequency components. Each point on the membrane resonates at a different frequency, and the spacing of these resonant frequencies along the BM is almost logarithmic. The effective frequency selectivity is related to the width of the filter characteristic at each point.

The *outer hair cells*, distributed along the length of the BM, react to feedback from the brainstem. They alter their length to change the resonant properties of the BM. As a consequence, the frequency response of the membrane becomes amplitude dependent. Finally, the inner hair cells of the basilar membrane fire when the BM moves upward. In doing so, they transduce the sound wave at each point into a signal on the auditory nerve. In this way the signal is half wave rectified. Each cell needs a certain time to recover between successive firings, so the average response during a steady tone is lower than at its onset. This means that the inner hair cells act as an automatic gain control.

The net result of the process described above is that an audio signal, which has a relatively wide-bandwidth, and large dynamic range, is encoded for transmission along the nerves. Each one of these nerves offers a much narrower bandwidth, and limited dynamic range. In addition, a critical process has happened during these steps. Any information that is lost due to the transduction

process within the cochlea is not available to the brain. In other words, the cochlea acts as a lossy coder. The vast majority of what we cannot hear is attributable to this transduction process (Robinson & Hawksford, 1999).

4.2 Perceptual Phenomena

As was just stated, one can model the processes that take place inside the HAS in order to represent how a listener responds to auditory stimuli. Given its characteristics, the HAS responds differently depending on the frequency and loudness of the input. This means that all components of a watermark may not be equally perceptible. Moreover, it also denotes the need of using a perceptual model to effectively measure the amount of distortion that is imposed on an audio signal when a mark is embedded. Given this fact, in this section the main processes that need to be included on a perceptual model are presented.

Sensitivity refers to the ear's response to direct stimuli. In experiments designed to measure sensitivity, listeners are presented with isolated stimuli and their perception of these stimuli is tested. For example, a common test consists of measuring the minimum sound intensity required to hear a particular frequency (Cox et al., 2002). The main characteristics measured for sensitivity are *frequency* and *loudness*.

The responses of the HAS are frequency dependent; variations in frequency are perceived as different tones. Tests show that the ear is most sensitive to frequencies around 3 kHz and that sensitivity declines at very low (20 Hz) and very high (20 kHz) frequencies. Regarding loudness, different tests have been performed to measure sensitivity. As a general result, one can state that the HAS is able to discern smaller changes when the average intensity is louder. In other words, the human ear is more sensitive to changes in louder signals than in quieter ones.

The second phenomenon that needs to be taken into account is *masking*. A signal that is clearly audible if presented alone can be completely inaudible in the presence of another signal, the masker. This effect is known as masking, and the masked signal is called the maskee. For example, a tone might become inaudible in the presence of a second tone at a nearby frequency that is louder. In other words, masking is a measure of a listener's response to one stimulus in the presence of another.

Two different kinds of masking can occur: simultaneous masking and temporal masking (Swanson, Zhu, Tewfik, & Boney, 1998). In simultaneous masking, both the masker and the maskee are presented at the same time and are quasi-stationary (ITU, 2001). In temporal masking, the masker and the maskee are presented at different times.

The third effect that has to be considered is *pooling*. When multiple frequencies are changed rather than just one, it is necessary to know how to combine the sensitivity and masking information for each frequency. Combining the perceptibilities of separate distortions gives a single estimate for the overall change in the work. This is known as pooling.

4.3 ABX Listening Test

Audio quality is usually evaluated by performing a listening test. In particular, the ABX listening test is commonly used when evaluating the quality of watermarked signals. Other tests for audio watermark quality evaluation, such as the one described in (Arnold & Schilz, 2002), follow a similar methodology as well. Given this, it becomes desirable to create an automatic model that predicts the response observed from a human listener in such a procedure.

In an ABX test the listener is presented with three different audio clips: selection A (non-watermarked audio), selection B (the watermarked audio) and X (either A or B), drawn at random. The listener is then asked to decide if selection X is equal to A or B. The number of correct answers is the basis to decide if the watermarked audio is perceptually different than the original audio and one will, therefore, declare the watermarking algorithm as perceptible. In the other case, if the watermarked audio is perceptually equal to the original audio, the watermarking algorithm will be declared as *transparent*, or imperceptible.

The ABX test is fully described in ITU Recommendation ITU-R BS.1116, and has been successfully used for subjective measurement of impaired audio signals. Normally only one attribute is used for quality evaluation. It is also defined that this attribute represents any and all detected differences between the original signal and the signal under test. It is known as *basic audio quality* (BAQ), and is calculated as the difference between the grade given to the impaired signal and the grade given to the original signal. Each one of these grades uses the five level impairment scale.

Although its results are highly reliable, there are many problems related to performing an ABX test for watermark quality evaluation. One of them is the subjective nature of the test, as the perception

conditions of the listener may vary with time. Another problem arises form the high costs associated with the test. These costs include the setup of audio equipment, construction of a noise-free listening room, and the costs of employing individuals with extraordinary acute hearing. Finally, the time required to perform extensive testing also poses a problem to this alternative.

Given these facts it becomes desirable to automate the ABX listening test, and incorporate it into a perceptual model of the HAS. If this is implemented, then the task measuring perceptibility can be fully automated and thus watermarking schemes can be effectively and thoroughly evaluated. Fortunately, several perceptual models for audio processing have been proposed. Specifically, in the field of audio coding, psychoacoustic models have been successfully implemented to evaluate the perceptual quality of coded audio. These models can be used as a baseline performance tool for measuring the perceptibility of audio watermarking schemes.

4.4 A Perceptual Model

A perceptual model used for evaluation of watermarked content must compare the quality of two different audio signals in a way that is similar to the ABX listening test. These two signals correspond to the original audio cover *A* and the watermarked audio file *A'*. An ideal system will receive both signals as an input, process them through an auditory model, and compare the representations given by this model (Thiede et al., 1998). Finally it will return a score for the watermarked file *A'*, in the five level impairment scale. More importantly, the result of such an objective test must be highly correlated with those achieved under a subjective listening test (ITU, 2001).

The auditory model used to process the input signals will have a similar structure to that of the HAS. In general terms, the response of each one of the components of the HAS is modeled by a series of filters. In particular, a synopsis of the models proposed in (Robinson & Hawksford, 1999), (Thiede & Kabot, 1996), (Thiede et al., 1998), and (ITU, 2001) is now presented.

The filtering performed by the pinna and ear canal is simulated by an FIR filter, which has been derived from experiments with a dummy head. More realistic approaches can use measurements from human subjects. After this prefiltering, the audio signal has to be converted to a basilar membrane representation. That is, the amplitude dependent response of the basilar membrane needs to be simulated. In order to do this, the first step consists of processing the input signal through a bank of amplitude dependant filters, each one adapted to the frequency response of a point on the basilar membrane. The center frequency of each filter should be linearly spaced on the Bark scale, a commonly used frequency scale. The actual number of filters to be used depends on the particular implementation. Other approaches might use a Fast Fourier Transform to decompose the signal, but this creates a tradeoff between temporal and spectral resolution (Thiede & Kabot, 1996).

At each point in the basilar membrane, its movement is transduced into an electrical signal by the hair cells. The firing of individual cells is pseudorandom, but when the individual signals are combined, the proper motion of the BM is derived. Simulating the individual response of each hair cell and combining these responses is a difficult task, so other practical solutions have to be applied. In particular, (Robinson & Hawksford, 1999) implements a solution based on calculating the half wave response of the cells, and then using a series of feedback loops to simulate the increased sensitivity of the inner hair cells to the onset of sounds. Other schemes might just convolve the signal with a spreading function, to simulate the dispersion of energy along the basilar membrane, and then convert the signal back to decibels (ITU, 2001). Independently of the method used, the basilar membrane representation is obtained at this point.

After a basilar membrane representation has been obtained for both the original audio signal *A*, and the watermarked audio signal *A'*, the perceived difference between the two has to be calculated. The difference between the signals at each frequency band has to be calculated, and then it must be determined at what level these differences will become audible for a human listener (Robinson & Hawksford, 1999). In the case of the ITU Recommendation ITU-R BS.1387, this task is done by calculating a series of model variables, such as excitation, modulation and loudness patterns, and using them as an input to an artificial neural network with one hidden layer (ITU, 2001). In the model proposed in (Robinson & Hawksford, 1999), this is done as a summation over time (over an interval of 20 ms) along with weighting of the signal and peak suppression.

The result of this process is an objective difference between the two signals. In the case of the ITU model, the result is given in a negative five level impairment scale, just like the BAQ, and is known as the Objective Difference Grade (ODG). For other models, the difference is given in implementation-dependant units. In both cases, a

mapping or scaling function, from the model units to the ITU-R. 500 scale, must be used.

For the ITU model, this mapping could be trivial, as all that is needed is to add a value of 5 to the value of the ODG. However, a more precise mapping function could be developed. The ODG has a resolution of one decimal, and the model was not specifically designed for the evaluation watermarking schemes. Given this, a non-linear mapping (for example using a logarithmic function), could be more appropriate.

For other systems, determining such a function will depend on the particular implementation of the auditory model; nonetheless such a function should exist, as a correlation between objective and subjective measures was stated as an initial requirement. For example, in the case of (Thiede & Kabot, 1996), a sigmoidal mapping function is used. Furthermore, the parameters for the mapping function can be calculated using a control group consisting of widely available listening test data.

The resulting grade, in the five level scale, is defined as the *perceptibility* of the audio watermark. This means that in order to estimate the perceptibility of the watermarking scheme, several test runs must be performed. Again, these test runs should embed a random mark on a cover signal, and a large and representative set of audio cover signals must be used. The perceptibility test score is finally calculated by averaging the different results obtained for each one of the individual tests.

5 FINAL BENCHMARK SCORE

In the previous sections three different testing procedures have been proposed, in order to measure the fidelity, robustness and perceptibility of a watermarking scheme. Each one of these tests has resulted in several scores, some of which may be more useful than others. These scores have to be combined in order to obtain a final benchmarking score. As a result, fair comparison amongst competing technologies can be possible, as the final watermarking scheme evaluation score is obtained.

In addition, another issue is addressed at this point: defining the specific parameters to be used for each attack while performing the robustness test. While the different attacks were explained previously, the strength at which they should be applied was not specified.

Addressing these two topics can prove to be a difficult task. Moreover, a single answer might not be appropriate for every possible watermarking application. Given this fact, one should develop and use a set of application-specific evaluation templates to overcome this restriction. In order to do so, an *evaluation template* is defined as a set of guidelines that specifies the specific parameters to be used for the different tests performed, and also denotes the relative importance of each one of the tests performed on the watermarking scheme. Two fundamental concepts have been incorporated into that of evaluation templates: evaluation profiles and application specific benchmarking.

Evaluation profiles have been proposed in (Petitcolas, 2000) as a method for testing different levels of robustness. Their sole purpose is to establish the set of tests and media to be used when evaluating a marking algorithm. For example, one should test a marking scheme intended for advertisement broadcast monitoring with a set of recordings similar to those that will be used in a real world situation. There is no point in testing such an algorithm with a set of high-fidelity musical recordings. Evaluation profiles are thus a part of the proposed evaluation templates.

Application specific benchmarking, in turn, is proposed in (Pereira et al., 2001; Voloshynovskiy, Pereira, Iquise, & Pun, 2001) and consists of averaging the results of the different tests performed to a marking scheme, using a set of weights that is specific to the intended application of the watermarking algorithm. In other words, attacks are weighted as a function of applications (Pereira et al., 2001). In the specific case of the evaluation templates proposed in this document, two different sets of weights should be specified: those used when measuring one of the three fundamental characteristics of the algorithm (i.e., fidelity, robustness and perceptibility); and those used when combining these measures into a single benchmarking score.

After the different weights have been established, the *overall watermarking scheme score* is calculated as a simple weighted average, with the formula

$$Score = w_f * s_f + w_r * s_r + w_p * s_p , \qquad (4)$$

where w represents the assigned weight for a test, s to the score received on a test, and the subscripts f, r, p denote the fidelity, robustness and perceptibility tests respectively. In turn, the values of s_f, s_r, and s_p are also determined using a weighted average for the different measures obtained on the specific subtests.

The use of an evaluation template is a simple, yet powerful idea. It allows for a fair comparison of watermarking schemes, and for ease of automated testing. After these templates have been

defined, one needs only to select the intended application of the watermarking scheme that is to be evaluated, and the rest of the operations can be performed automatically. Nonetheless, time has to be devoted to the task of carefully defining the set of evaluation templates for the different applications sought to be tested.

5.1 Presenting the Results

The main result of the benchmark presented here is the overall watermarking scheme score that has just been explained. It corresponds to a single, numerical result. As a consequence, comparison between similar schemes is both quick and easy. Having such a comprehensive quality measure is sufficient in most cases.

Under some circumstances the intermediate scores might also be important, as one might want to know more about the particular characteristics of a watermarking algorithm, rather than compare it against others in a general way. For these cases, the use of graphs, as proposed in (Kutter & Hartung, 2000; Kutter & Petitcolas, 1999, 2000) is recommended.

The graphs should plot the variance in two different parameters, with the remaining parameters fixed. That is, the test setup conditions should remain constant along different test runs. Finally, several test runs should be performed, and the results averaged. As a consequence, a set of variable and fixed parameters for performing the comparisons are possible, and thus several graphs can be plotted. Some of the most useful graphs for this task are presented in (Kutter & Petitcolas, 1999), along with their corresponding variables and constants.

6 CONCLUSION

The watermarking benchmark proposed here can be implemented for the automated evaluation of different watermarking schemes. In fact, this idea has been included in test design, and has motivated some key decisions, such as the use of a computational model of the ear instead of a formal listening test. Moreover, the establishment of an automated test for watermarking systems is an industry need, as tird-party evaluation of watermarking schemes seems to be the only objective solution to the problem transparent evaluation (Petitcolas, 2000).

As a conclusion, the industry needs to establish a trusted evaluation authority in order to objectively evaluate its watermarking products. The establishment of watermark certification programs has been proposed, and projects such as the Certimark and StirMark benchmarks are under development (Certimark, 2001; Kutter & Petitcolas, 2000; Pereira et al., 2001; Petitcolas et al., 2001). However, these programs seem to be aimed mainly at testing of image watermarking systems (Meerwald & Pereira, 2002). A similar initiative for audio watermark testing has yet to be proposed.

REFERENCES

Acken, J. M. (1998, July 1998). How watermarking adds value to digital content. *Communications of the ACM, 41,* 75-77.

Arnold, M. (2000). *Audio watermarking: Features, applications and algorithms.* Paper presented at the IEEE International Conference on Multimedia and Expo 2000.

Arnold, M., & Schilz, K. (2002, January 2002). *Quality evaluation of watermarked audio tracks.* Paper presented at the Proceedings of the SPIE, Security and Watermarking of Multimedia Contents IV, San Jose, CA.

Certimark. (2001). *Certimark benchmark, metrics & parameters* (D22). Geneva, Switzerland.

Cox, I. J., Miller, M. L., & Bloom, J. A. (2002). *Digital Watermarking* (1 ed.). San Francisco: Morgan Kaufmann.

Craver, S., Memon, N., Yeo, B.-L., & Yeung, M. M. (1998). Resolving rightful ownerships with invisible watermarking techniques: Limitations, attacks and implications. *IEEE Journal on Selected Areas in Communications, 16*(4), 573-586.

Craver, S., Memon, N., & Yeung, M. M. (1996). *Can invisible watermarks resolve rightful ownerships?* (RC 20509): IBM Research.

Craver, S., Perrig, A., & Petitcolas, F. A. P. (2000). Robustness of copyright marking systems. In F. A. P. Petitcolas & S. Katzenbeisser (Eds.), *Information hiding: Techniques for steganography and digital watermarking* (1 ed., pp. 149-174). Boston: Artech House.

Craver, S., Yeo, B.-L., & Yeung, M. M. (1998, July 1998). Technical trials and legal tribulations. *Communications of the ACM, 41,* 45-54.

Dannenberg, R., & Mazzoni, D. (2002). Audacity (Version 0.98). Pittsburgh, PA.

Gordy, J. D., & Burton, L. T. (2000, August 2000). *Performance evaluation of digital audio water-marking algorithms.* Paper presented at the 43rd Midwest Symposium on Circuits and Systems, Lansing, MI.

ITU. (2001). *Method for objective measurements of perceived audio quality* (ITU-R BS.1387). Geneva: International Telecommunication Union.

Johnson, N. F., & Katzenbeisser, S. C. (2000). A survey of steganographic techniques. In F. A. P. Petitcolas & S. Katzenbeisser (Eds.), *Information hiding: Techniques for steganography and digital watermarking* (1 ed., pp. 43-78). Boston: Artech House.

Katzenbeisser, S., & Veith, H. (2002, January 2002). *Securing symmetric watermarking schemes against protocol attacks.* Paper presented at the Proceedings of the SPIE, Security and Watermarking of Multimedia Contents IV, San Jose, CA.

Kirovski, D., & Malvar, H. (2001, April 2001). *Robust cover communication over a public audio channel using spread spectrum.* Paper presented at the Information Hiding Workshop, Pittsburgh, PA.

Kutter, M., & Hartung, F. (2000). Introduction to watermarking techniques. In F. A. P. Petitcolas & S. Katzenbeisser (Eds.), *Information hiding: Techniques for steganography and digital watermarking* (1 ed., pp. 97-120). Boston: Artech House.

Kutter, M., & Petitcolas, F. A. P. (1999, January 1999). *A fair benchmark for image watermarking systems.* Paper presented at the Electronic Imaging '99. Security and Watermarking of Multimedia Contents, San Jose, CA.

Kutter, M., & Petitcolas, F. A. P. (2000). Fair evaluation methods for image watermarking systems. *Journal of Electronic Imaging, 9*(4), 445-455.

Meerwald, P., & Pereira, S. (2002, January 2002). *Attacks, applications, and evaluation of known watermarking algorithms with Checkmark.* Paper presented at the Proceedings of the SPIE, Security and Watermarking of Multimedia Contents IV, San Jose, CA.

Pereira, S., Voloshynovskiy, S., Madueño, M., Marchand-Maillet, S., & Pun, T. (2001, April, 2001). *Second generation benchmarking and application oriented evaluation.* Paper presented at the Information Hiding Workshop, Pittsburgh, PA.

Petitcolas, F. A. P. (2000). Watermarking schemes evaluation. *IEEE Signal Processing, 17*(5), 58-64.

Petitcolas, F. A. P., & Anderson, R. J. (1998, September 1998). *Weaknesses of copyright marking systems.* Paper presented at the Multimedia and Security Workshop at the 6th ACM International Multimedia Conference, Bristol U.K.

Petitcolas, F. A. P., & Anderson, R. J. (1999, June, 1999). *Evaluation of copyright marking systems.* Paper presented at the IEEE Multimedia Systems, Florence, Italy.

Petitcolas, F. A. P., Anderson, R. J., & G., K. M. (1998, April 1998). *Attacks on copyright marking systems.* Paper presented at the Second workshop on information hiding, Portland, OR.

Petitcolas, F. A. P., Steinebach, M., Raynal, F., Dittmann, J., Fontaine, C., & Fatès, N. (2001, January 22-26). *A public automated web-based evaluation service for watermarking schemes: StirMark Benchmark.* Paper presented at the Electronic Imaging 2001, Security and Watermarking of Multimedia Contents, San Jose, CA.

Piron, L., Arnold, M., Kutter, M., Funk, W., Boucqueau, J. M., & Craven, F. (1999, January, 1999). *OCTALIS benchmarking : Comparison of four watermarking techniques.* Paper presented at the Proceedings of SPIE: Security and Watermarking of Multimedia Contents, San Jose, CA.

Robinson, D. J. M., & Hawksford, M. J. (1999, September 1999). *Time-domain auditory model for the assessment of high-quality coded audio.* Paper presented at the 107th Conference f the Audio Engineering Society, New York, NY.

Swanson, M. D., Zhu, B., Tewfik, A. H., & Boney, L. (1998). Robust audio watermarking using perceptual masking. *Signal Processing, 66*(3), 337-355.

Thiede, T., & Kabot, E. (1996, 1996). *A new perceptual quality measure for bit rate reduced audio.* Paper presented at the 100th AES Convention, Copenhagen, Denmark.

Thiede, T., Treurniet, W. C., Bitto, R., Sporer, T., Brandenburg, K., Schmidmer, C., Keyhl, K., G., B. J., Colomes, C., Stoll, G., & Feiten, B. (1998, 1999). *PEAQ - der künftige ITU-Standard zur objektiven messung der wahrgenommenen audioqualität.* Paper presented at the Tonmeistertagung Karlsruhe, Munich, Germany.

Voloshynovskiy, S., Pereira, S., Iquise, V., & Pun, T. (2001, June 2001). *Attack modelling: towards a second generation benchmark.* Paper presented at the Signal Processing.

Voloshynovskiy, S., Pereira, S., Pun, T., Eggers, J. J., & Su, J. K. (2001, August, 2001). Attacks on digital watermarks: Classification, estimation-based attacks and benchmarks. *IEEE Communications Magazine, 39,* 118-127.

COMPRESSION OF HYPERSPECTRAL IMAGERY VIA LINEAR PREDICTION

Francesco Rizzo, Bruno Carpentieri
Dipartimento di Informatica ed Applicazioni "R.M. Capocelli", Università degli Studi di Salerno, Via S. Allende, Baronissi (SA), Italy
Email: frariz@ieee.org, bc@dia.unisa.it

Giovanni Motta, James A. Storer
Computer Science Department, Brandeis University, Waltham 02454 MA, USA
Email: gim@ieee.org, storer@cs.brandeis.edu

Keywords: Predictive Coding, Data Compression, Remote Sensing, 3D Data.

Abstract: (Motta et al., 2003) proposed a Locally Optimal Vector Quantizer (LPVQ) for lossless encoding of hyperspectral data, in particular, Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) images. In this paper we first show how it is possible to improve the baseline LPVQ algorithm via linear prediction techniques, band reordering and least squares optimization. Then, we use this knowledge to devise a new lossless compression method for AVIRIS images. This method is based on a low complexity, linear prediction approach that exploits the linear nature of the correlation existing between adjacent bands. A simple heuristic is used to detect contexts in which such prediction is likely to perform poorly, thus improving overall compression and requiring only marginal extra storage space. A context modeling mechanism coupled with a one band look ahead capability allows the proposed algorithm to match LPVQ compression performances at a fraction of its space and time requirements. This makes the proposed method suitable to applications where limited hardware is a key requirement, spacecraft on board implementation. We also present a least squares optimized linear prediction for AVIRIS images which, to the best of our knowledge, outperforms any other method published so far.

1 INTRODUCTION

In the last three decades, air-borne and space-borne remote acquisition of high definition electro-optic images has been increasingly used in military and civilian applications to recognize objects and classify materials on the earth's surface. By analyzing the spectrum of the reflected light it is possible to recognize the material(s) composing the observed scene. The development of new detector technologies has made possible the introduction of new classes of aircraft spectrometers capable of recording a large number of spectral bands over the visible and reflected infrared region. For this reason the data sets they produce are often referred to as hyperspectral. These instruments have reached spectral resolution sufficient to allow very accurate characterization of the spectral reflectance curve of a given spatial area. For example, images acquired with the JPL's Airborne Visible/Infrared Imaging Spectrometer, AVIRIS (NASA, 2003), have pixels covering an area of approximately 20x20 meters, with radiance decomposed into 224 narrow bands, approximately 10*nm* wide each, in the range 400-2,500*nm*. Spectral components are represented with a 16 bits precision.

Hyperspectral imagery is a rapidly growing source of remote sensed data, even though its precision pales compared to the millions of channels of a truly high resolution lab spectrometer. The technology seems mature enough to use higher resolution, space-borne spectrometers. In fact, increasing the number of bands, i.e. the spectral resolution, allows for more sophisticated analysis and increases the data rate by only a linear amount. The problem is that the acquisition of these images already produces large amounts of highly correlated

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 275–282.

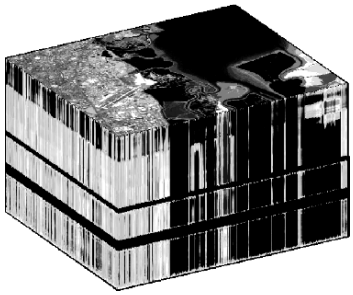


Figure 1: AVIRIS data cube Moffett Field, scene 1 (NASA, 2003).

data (e.g., in the range 140-1,000 Mb for AVIRIS images) in the form of a two dimensional image matrix each pixel consisting of many components, one for each spectral band (Figure 1).

Since hyperspectral imagery is acquired at cost and often used in critical tasks like classification (assignment of a label to every pixel) or target detection (identification of a somewhat rare instance), compression algorithms that provide lossless or near-lossless quality (for classification and detection purposes) may be required. In addition, it may be desirable to have low complexity that allows efficient on-board implementation with limited hardware. Traditional approaches to the compression of hyperspectral imagery are based on differential prediction via DPCM (Aiazzi, 2001; Abousleman, 1995; Abousleman et al., 2002), direct vector quantization (Manohar and Tilton, 2000; Ryan and Arnold, 1997; Mielikäinen and Toivanen, 2002; Pickering and Ryan, 2001) or dimensionality reduction through Principal Component Analysis.

In (Motta et al., 2003) a locally optimal design of a partitioned vector quantizer (LPVQ) for the encoding of high dimensional data is presented. The algorithm is applied to lossless, near-lossless and lossy compression of AVIRIS data. LPVQ's lossless compression, is aligned with the current state of the art. Its design and coding process, on the other hand, are computationally intensive (although highly parallelizable), while decoding is just table lookup. The asymmetrical nature of the algorithm makes it most appropriate for systems in which the codebook design does not have to be performed on-board. An inter-band linear prediction approach based on least square optimization is presented in (Mielikäinen et al., 2002). This compression method optimizes the parameters of a linear predictor with spatial and spectral support. Such optimization is performed for each sample.

Using linear prediction, least square optimization, and optimal band reordering, in Section 2 we show how to encode efficiently the quantization indices produced by LPVQ, improving upon the baseline algorithm. We also exploit successfully the fact that spectral correlation in the original data is preserved in LPVQ indices after quantization.

In Section 3 we target the linear nature of the spectral correlation of AVIRIS data with a simple linear prediction method. The proposed method is composed by an *intra-band* predictor, similar to the one in LOCO-I (Weinberger et al., 2000), for the few bands with strong spatial correlation. The rest is encoded using a novel *inter-band* predictor. This predictor shares the same low complexity of the intra-band one, and requires buffering of at most two scan-lines from each of the previous three bands. It also uses a simplified version of the context modeling mechanism in LOCO-I that allows to mach the compression performance of LPVQ. Finally we discuss experimental results and current research directions.

2 IMPROVING ENTROPY CODING OF LPVQ'S QUANTIZATION INDICES

(Motta et al., 2003) compress hyperspectral data by using a modified version of the Generalized Lloyd Algorithm to perform a dimensionality reduction of the original data. The *D*-dimensional input vectors are broken into *L* sub–vectors (*L=16* in the reported experiments). Each sub–vector is then encoded with the 8–bit index of the closest match in the codebook generated by LPVQ, while the quantization error is encoded separately. The spatial correlation in the original data is preserved in the index files (planes), so they look very much like "natural" grayscale images. The index files are then encoded using LOCO-I.

In this section we focus on improving the compression of the quantization indices. We note that spectral dependency is still observable among index files. To take advantage of this phenomenon, we propose three methods (summarized in Table 1) two of which extend the LOCO-I/JPEG-LS predictor. They compute the prediction, based on a causal data subset (Figure 2), $\hat{x}_{i,j,k}$ of the pixel $x_{i,j,k}$ in the *i*-th row, *j*-th column of the *k*-th plane.

The first method in Table 1 is the one used by LOCO-I, reported here as a reference. The second, that we call INTER predictor, is similar to the one presented in (Barequet an Feder, 1999), while 3D-MED is a novel, general extension of LOCO-I to an inter-band context. These two methods share the

Table 1: Linear predictors for encoding of LPVQ quantization indices.

$\hat{x}_{\text{LOCO-I}} = \text{Median}(x_{i,j-1,k}, x_{i-1,j,k}, x_{i,j-1,k} + x_{i-1,j,k} - x_{i-1,j-1,k})$
$\hat{x}_{\text{INTER}} = x_{i,j,k-1} + \text{Median}(D_{1,k}, D_{2,k}, D_{1,k} + D_{2,k} - D_{3,k})$ $D_{1,k} = x_{i,j-1,k} - x_{i,j-1,k-1},\ D_{2,k} = x_{i-1,j,k} - x_{i-1,j,k-1},\ D_{3,k} = x_{i-1,j-1,k} - x_{i-1,j-1,k-1}$
$\hat{x}_{\text{3D-MED}} = \text{Median}(x_{i,j-1,k}, x_{i-1,j,k}, x_{i,j,k-1}, x_{i,j-1,k} + x_{i-1,j,k} - x_{i-1,j-1,k}, x_{i,j-1,k} + x_{i,j,k-1} - x_{i,j-1,k-1},$ $x_{i-1,j,k} + x_{i,j,k-1} - x_{i-1,j,k-1})$
$\hat{x}_{\text{3D-LSQ}} = \sum_{j=1}^{N} \alpha_j x_j$

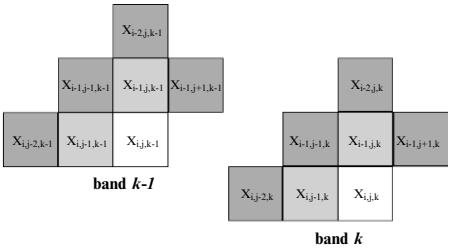


Figure 2: Prediction Template.

same low complexity of the JPEG-LS standard, and hence a highly efficient implementation is possible.

The third method, 3D-LSQ, is more aggressive and computationally more expensive: given a reference plane and a 3D subset of causal data, an optimal linear predictor, in the least square sense, is determined for each sample. The prediction structure and the notation used in the following is similar to the one presented in (Brunello et al., 2002).

Two different context enumerations are defined based on the distance functions

$$d_{2D}(x_{m,n,k}, x_{p,q,k}) = \sqrt{(m-p)^2 + (n-q)^2}$$

$$d_{3D}(x_{m,n,i}, x_{p,q,j}) = \begin{cases} \sqrt{(m-p)^2 + (n-q)^2} & j = i \\ \sqrt{\frac{1}{4} + (m-p)^2 + (n-q)^2} & j \neq i \end{cases}$$

The resulting 2D and 3D context templates are showed in Figure 3.

In the following, by $x(i)$ we denote the i-th pixel in the above enumeration of the 2D context of $x_{i,j,k}$. Moreover, $x(i, j)$ denotes the j-th pixel in the 3D context of $x(i)$. The N-th order prediction of the current pixel ($x_{m,n,k} \equiv x(0,0)$, we drop the subscript and the parenthesis when referring to the current pixel) is computed as

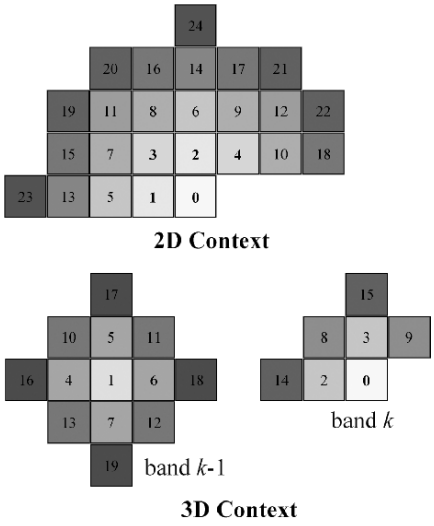


Figure 3: 2D and 3D contexts and pixel enumerations.

$$\hat{x}(0,0) = \sum_{j=1}^{N} \alpha_j \cdot x(0, j)$$

The coefficients $\alpha_0 = [\alpha_1, \text{K}\ , \alpha_N]^t$ minimizing the energy of the prediction error

$$P = \sum_{i=1}^{M} \left(x(i,0) - \hat{x}(i,0)\right)^2$$

are calculated using the well-known theory on optimal linear prediction. Notice that the data used in the prediction are a causal, finite sub–set of the past data and no side information needs to be sent to the decoder.

Using matrix notation, we write

$$P = (\mathbf{C}\alpha - \mathbf{X})^t \cdot (\mathbf{C}\alpha - \mathbf{X})$$

where,

$$\mathbf{C} = \begin{bmatrix} x(1,1) & \Lambda & x(1,N) \\ \mathrm{M} & \mathrm{O} & \mathrm{M} \\ x(M,1) & \Lambda & x(M,N) \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} x(1,0) \\ \mathrm{M} \\ x(M,0) \end{bmatrix}$$

By taking the derivative with respect to α and setting it to zero, the optimal predictor coefficients are the solution of the following linear system

$$(\mathbf{C}'\mathbf{C}) \cdot \alpha_0 = \mathbf{C}'\mathbf{X}$$

Once the optimal predictor coefficients for the current sample have been determined, the prediction error $\varepsilon = \lfloor x - \hat{x} \rfloor$ is encoded in the same way of the previous two methods.

3 OPTIMAL BAND ORDERING

Because each index file represents a subset of contiguous spectral bands, and because the correlation between two bands is not always inversely proportional to the distance of their wavelengths, a sequential encoding of the index files is generally suboptimal. In order to address this issue, given a function $f(i, j)$ representing the cost of encoding plane *j* using plane *i* as reference, it is possible to find the optimal plane ordering using standard graph theory results. Similar ideas could be find in (Tate, 1997; Motta and Weinberger, 2001).

Given the cost $f(i, j)$, we can define a complete weighted graph with *L* nodes where the weight of the edge $w_{i,j}$ is equal to $f(i, j)$. We add a fictitious node 0 connected by an edge to each node *j*. The weight $w_{\mathbf{0},j}$ represents the cost of encoding plane j without using any reference plane (e.g., using LOCO-I). The problem of optimal plane ordering is equivalent to the problem of finding the minimum spanning tree of the resulting graph (if $f(i, j)$ is not symmetrical then the graph is directed and one should compute the optimal branching rooted at **0** (Gabow et al., 1986)).

As a proof of concept, we used the first order entropy of the difference between each pair of planes as a cost function.

3.1 Context Modeling

The underlying assumption of the previous section is that the index planes generated by LPVQ look very much as "natural" images. This justifies the use of off-shelf image-oriented techniques to encode these data. This behavior is the by-product of the lexicographical sorting of the centroids generated by LPVQ, which are "scaled/translated" version of each other. Similar behaviors are experienced in standard VQ image compression when code-vectors are arranged by increasing norm. This is not surprising because if the VQ is not rate-distortion optimal (like in most practical applications), then there must exist some inter-codeword correlation. Given the structure of LPVQ, there must be some correlation between the codeword of adjacent sub-vectors as well, hence the previous assumption is sub-optimal.

A lossless block coding of VQ code-vectors specifically designed for image compression, Address-VQ, was proposed in (Nasrabadi and Feng, 1990). Improvements were presented in (Wu et al., 1998; Gong et al., 2000), which exploited the inter-codeword correlations by means of context modeling and conditional entropy. These methods are *off-line* algorithms based on Bayes' theorem

$$P(X \mid X_1, X_2) = \frac{P(X_1, X_2 \mid X)}{P(X_1, X_2)}$$

where X is the VQ index to be coded, X_1 and X_2 causal neighbor of X).

The 3-dimensional nature of the LPVQ index planes suggests the use of a 3-D causal context. In order to assess the potentials of a Bayesian context modeling scheme, we analyzed the empirical probability $P(x_{i,j,k} \mid x_{i,j,k-1}, x_{i,j-1,k} - x_{i-1,j,k})$, and $P(x_{i,j,k} \mid x_{i,j,k-1})$ for each index plane. In general, the value of the pixel in the current plane is better predicted by the value of the corresponding pixel in the previous plane. This suggests a very simple, *on-line* scheme named PREV: define 256 entropy coders; encode $x_{i,j,k}$ using the $x_{i,j,k-1}$-th coder (without any form of prediction).

```
proc PREV
     def EC[256], EC1 as entropy_coder

     ; encode x_{i,j,1} using EC1
     EncodePlane(1,EC1)

     for K = 2 to L do
        for I = 1 to ROWS do
           for J = 1 to COLS do
              Encode(x_{i,j,k},EC[x_{i,j,k-1}])
end proc
```

Table 2: Entropy coding results for LPVQ indices.

	LOCO	Sequential Coding			Optimal Band Ordering			PREV
		INTER	3D-MED	3D-LSQ	INTER	3D-MED	3D-LSQ	
Cuprite	40.44	43.44 +7.42%	44.97 +11.19%	48.82 +20.73%	47.30 +16.96%	46.11 +14.01%	50.28 +24.33%	50.52 +24.93%
Jasper Ridge	35.02	35.99 +2.77%	37.75 +7.82%	39.13 +11.76%	38.06 +8.68%	38.39 +9.64%	39.88 +13.89%	47.31 +35.09%
Low Altitude	39.10	40.61 +3.86%	42.10 +7.67%	45.96 +17.54%	44.33 +13.38%	43.15 +10.36%	47.12 +20.52%	51.93 +32.78%
LunarLake	44.45	46.30 +4.15%	48.00 +7.99%	51.22 +15.23%	49.42 +11.18%	48.91 +10.04%	52.38 +17.85%	57.72 +29.85%
Moffet Field	40.92	43.35 +5.92%	44.28 +8.19%	47.34 +15.67%	46.53 +13.70%	45.28 +10.65%	48.76 +19.14%	55.70 +36.12%
AVERAGE	**39.99**	**41.94** +4.87%	**43.42** +8.59%	**46.49** +16.28%	**45.13** +12.86%	**44.37** +10.96%	**47.68** +19.25%	**52.64** +31.63%

3.2 Experimental results

Table 2 reports results in terms of compression far all schemes presented so far. The reported results for 3D-LSQ are obtained with $M = 90$ and $N = 9$.

As expected, when sequential coding is used, 3D-LSQ is better than the 3D-MED and the INTER predictor. Compared to the baseline LOCO-I coding, on average the improvements attained are respectively +16%, +8.54% and +4.87. When the optimal plane ordering is in use, the improvements are much higher (+19.25% for 3D-LSQ). More interestingly, the PREV prediction/compression scheme is more than 10% better than 3D-LSQ with optimal ordering, and more than 30% better than LOCO-I, used in (Motta et al., 2003). Furthermore, PREV is more than 200 times faster than 3D-LSQ on a AMD Athlon(tm) MP 1900+ based personal computer.

4 INTER-BAND LINEAR PREDICTION

Remote sensed images, like AVIRIS, show two forms of correlation: spatial (the same material tends to be present in many adjacent pixels: e.g., the water of a river) and spectral (one band can be fully or partially predicted from other bands). From our investigations emerges that the spectral correlation is generally much stronger than the spatial correlation. Furthermore, dynamic range and noise levels (instrument noise, reflection interference, aircraft movements, etc.) of AVIRIS data are much higher than those in photographic images. For these reasons the spatial predictor of LOCO-I (Table 1) tends to fail on this kind of data. Figure 4 shows the performance in terms of bit per sample of this

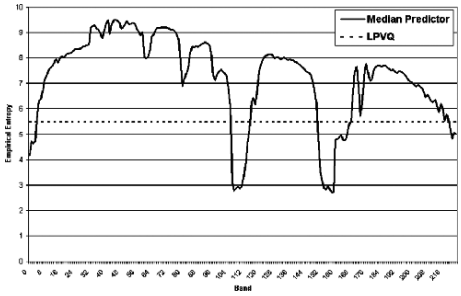


Figure 4: Empirical band entropy of the *Median* predictor.

predictor. From our simulations it is clear that the median predictor of JPEG-LS is inefficient almost everywhere, and especially in the visible part of the spectrum that accounts for almost half of the data and it is characterized by large dynamic ranges. Nevertheless, JPEG-LS fast and efficient compression would be highly desirable to an on-board, hardware implementation.

Motivated by these considerations, we propose a novel compression method for AVIRIS data using a a novel predictor for bands marked *inter-band* (*IB* set) and a linear predictor in the style of JPEG-LS for the rest.

This new predictor uses a simple heuristic to detect contexts in which it is likely to fail. In such cases the prediction is corrected using information about the behavior of the inter-band predictor in the previous two bands. After this prediction step, the prediction error is computed and entropy coded with a simple arithmetic coder. See Figure 5 for a formal description. After the prediction step, the prediction error is computed and entropy coded with a simple arithmetic coder.

$$\hat{x}_{i,j,k} = \begin{cases} f_{i,j,k} = x_{i,j,k-1} + \left\lfloor \dfrac{D_{1,k} + D_{2,k} + D_{3,k}}{3} \right\rfloor & k \in IB,\ \max\limits_{d=1,2,3}(D_{d,k}) - \min\limits_{d=1,2,3}(D_{d,k}) < T \\ f_{i,j,k} + \left\lfloor \dfrac{(x_{i,j,k-1} - f_{i,j,k-1}) + (x_{i,j,k-1} - f_{i,j,k-2})}{2} \right\rfloor & k \in IB,\ \max\limits_{d=1,2,3}(D_{d,k}) - \min\limits_{d=1,2,3}(D_{d,k}) \geq T \\ \hat{x}_{LOCO-I}(i,j,k) & k \notin IB \end{cases}$$

Figure 5: Inter-Band Linear Predictor.

4.1 Least Squares Optimization

In order to set an upper bound for the achievable compression by the proposed linear prediction method and for the data under examination, we decided to implement a prediction scheme optimized for each pixel and for each band based on least squares optimization. We apply the 3D-LSQ (here named SLSQ) approach of Section 2 directly to the 224 AVIRIS bands, rather than the 16 LPVQ index planes, with $M = 4$ and $N = 1$.

The lossless compression results achieved by this method on AVIRIS images are, at the best of our knowledge, better than those published so far.

4.2 Experimental Results

Table 3 reports the compression ratio obtained by LP and SLSQ on the five "standard" publicly available AVIRIS images. We compare it with JPEG-LS, JPEG2000 (Taubman and Marcellin, 2001), and LPVQ. We do not report the compression results of (Mielikäinen et al., 2002) (claiming average compression ratio of 3.06:1). This is because their experimental results refer to a data sets that seems to be a subset of the one we are using and that we do not have currently available (furthermore (Mielikäinen et al., 2002) reports non-standard dimensions for AVIRIS images). LP has been applied with $IB = \Sigma\text{-}\{1\ldots8\}$, where Σ is the set of bands, and no prediction threshold. The proposed LP method is comparable to LPVQ at a fraction of the computational cost and it is sensibly superior to the standard lossless image coders.

We also tested an extension of LP and SLSQ based on the considerations taken from Figure 4. For each scene of each cube (28 total) we checked which band was better compressed spatially (LOCO-I) rather than spectrally (LD/SLSQ). For any given band i, $i \in IB$ if and only if it has been compressed in intra mode more than 15 times over 28 (*HEU* option). A more aggressive approach (*OPT*) assumes that the encoder checks for the best method first. This requires virtually no side information (1 bit/band) and a one band look-ahead capability. For LP we also introduced a simplified version of the context modeling mechanism described in (Weinberger et al., 2000), named *LP-CTX*.

Results of improved algorithms are reported in Table 4. We report also results of *differential* JPEG-LS and *differential* JPEG2000, where by "differential" we mean that the previous band is subtracted from the current one for spectral decorrelation before applying JPEG-LS or JPEG2000. This pre-preprocessing steps improves the two standard algorithms by 40% and 53% respectively, but better compression is achieved by LP and SLSQ. As we can see, the LP-CTX with on band look-ahead improves by more than 2% the LP method, matching LPVQ compression performance at a cost of a small increase of storage requirements over baseline LP, while being 5% better than differential JPEG-LS/JPEG2000. Finally, SLSQ-OPT achieves the overall best compression. While this method needs a one band look-ahead, it has the advantage of requiring virtually no side information (1 bit/band), and since inter and intra mode could be performed in parallel, compression time is practically unchanged.

5 CONCLUSIONS

In the first part of this paper we present and analyze three linear prediction schemes for the encoding of the index planes generated by the LPVQ algorithm. The best method achieves ≈20% improvement upon the basic schemes presented in (Motta et al., 2003). In the final subsection we show that the assumption that the index planes are comparable to "natural" images is not completely true. We also show how a very simple context modeling can achieve even better compression.

In the second part of the paper we propose a novel approach for lossless coding of AVIRIS data. It is based on an inter-band, linear predictor that, coupled with a simple entropy coder, competes with the current state of the art. The low complexity of the proposed method and its raster scan nature, makes it amenable for on-board implementations.

Since the proposed method depends loosely on the entropy coder, it would be also possible to remove the arithmetic coder and use the CCSDS standard algorithm for lossless data compression for space applications (CCSDS, 1997), whose hardware implementation is widely used on many satellites.

Table 3: Compression Results.

	JPEG-LS	JPEG2000	LPVQ	LPVQ-PREV	LP	SLSQ
Cuprite	2.09	1.91	3.13	3.18	3.03	3.15
Jaspder Ridge	2.00	1.80	2.82	2.88	2.94	3.15
Low Altitude	2.14	1.96	2.89	2.94	2.76	2.98
Lunar Lake	1.99	1.82	3.23	3.28	3.05	3.15
Moffett Field	1.91	1.78	2.94	3.00	2.88	3.14
AVERAGE	**2.03**	**1.85**	**3.00**	**3.06**	**2.93**	**3.12**

Table 4: Improvements of baseline LP and SLSQ algorithms.

	Differential JPEG-LS	Differential JPEG2000	LP-CTX			SLSQ	
			$IB = \Sigma\text{-}\{1\ldots8\}$	60%	OPT	60%	OPT
Cuprite	2.91	2.92	3.04	3.07	3.09	3.23	3.24
Jasper Ridge	2.81	2.82	2.96	2.98	3.00	3.22	3.23
Low Altitude	2.70	2.69	2.79	2.79	2.83	3.02	3.04
Lunar Lake	2.93	2.94	3.06	3.08	3.10	3.23	3.23
Moffett Field	2.84	2.83	2.93	2.94	2.96	3.20	3.21

We are currently working to improve the inter-band predictor and perform a formal analysis of the remaining correlation after prediction, in order to find suitable context modeling mechanisms that will indubitably improve current performances. Near-lossless extensions are also under consideration.

REFERENCES

Abousleman, G. P. (1995). Compression of hyperspectral imagery using hybrid DPCM/DCT and entropy constrained trellis coded quantization. In Storer, J. A. and Cohn, M., editors, *Proceedings Data Compression Conference*, pages 322–331. IEEE Computer Society Press.

Abousleman, G. P., Lam, T.-T., and Karam, L. J. (2002). Robust hyperspectral image coding with channeloptimized trellis-coded quantization. *IEEE Transactions on Geoscience and Remote Sensing*, 40(4):820–830.

Aiazzi, B., Alparone, L., and Baronti, S. (2001). Nearlossless compression of 3-D optical data. *IEEE Transactions on Geoscience and Remote Sensing*, 39(11):2547–2557.

Barequet, R. and Feder, M. (1999). SICLIC: A simple intercolor lossless image coder. In Storer, J. A. and Cohn, M., editors, *Proceedings of the Data Compression Conference*, pages 501–510, Snowbird, Utha. IEEE Computer Society Press.

Brunello, D., Calvagno, G., Mian, G. A., and Rinaldo, R. (2002). Lossless video coding using optimal 3D prediction. In *Proceedings of the 9th IEEE International Conference on Image Processing (ICIP 2002)*, volume 1, pages 89–92, Rochester, NY. IEEE Signal Processing Society.

CCSDS (1997). Consulting Committee for Space Data Systems, "Recommendation for space data system standards: Lossless data compression". CCSDS 121.0-B-1, Blue Book.

Gabow, H. N., Galil, Z., Spencer, T., and Tarjan, T. R. (1986). Efficient algorithms for finding minimum spanning trees in undirected and directed graphs. *Combinatorica*, 6(2):109–122.

Gong, Y., Fan, M. K. H., and Huang, C.-M. (2000). Image compression using lossless coding on VQ indexes. In Storer, J. A. and Cohn, M., editors, *Proceedings of the Data Compression Conference*, page 583, Snowbird, Utha. IEEE Computer Society Press.

Manohar, M. and Tilton, J. C. (2000). Browse level compression of AVIRIS data using vector quantization on massively parallel machine. In *Proceedings AVIRIS Airborne Geoscience Workshop*.

Mielikäinen, J., Kaarna, A., and Toivanen, P. (2002). Lossless hyperspectral image compression via linear prediction. *Proceedings of SPIE*, 4725(8):600–608.

Mielikäinen, J. and Toivanen, P. (2002). Improved vector quantization for lossless compression of AVIRIS images. In *Proceedings of the XI European Signal Processing Conference, EUSIPCO-2002*, Toulouse, France. EURASIP.

Motta, G., Rizzo, F., and Storer, J. A. (2003). On the compression of hyperspectral imagery. In Storer, J. A. and Cohn, M., editors, *Proceedings of the Data Compression Conference*, Snowbird, Utha. IEEE Computer Society Press.

Motta, G. and Weinberger, M. J. (2001). Compression of polynomial texture maps. Technical Report 143 (R.2), HP Laboratories Palo Alto.

NASA (2003). AVIRIS home page. http://popo.jpl.nasa.gov.

Nasrabadi, N. M. and Feng, Y. (1990). Image compression using address-vector quantization. *IEEE Transactions on Communication*, 38:2166–2173.

Pickering, M. and Ryan, M. (2001). Efficient spatialspectral compression of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7):1536–1539.

Ryan, M. J. and Arnold, J. F. (1997). The lossless compression of AVIRIS images by vector quantization. *IEEE Transactions on Geoscience and Remote Sensing*, 35(3):546–550.

Tate, S. R. (1997). Band ordering in lossless compression of multispectral images. *IEEE Transactions on Computers*, 46:477–483.

Taubman, D. and Marcellin, M. W. (2001). *Jpeg2000: Image Compression Fundamentals, Standards, and Practice*. Kluwer Academic Publishers, Boston, MA.

Weinberger, M. J., Seroussi, G., and Sapiro, G. (1996). LOCO-I: A low complexity, context-based, lossless image compression algorithm. In Storer, J. A. and Cohn, M., editors, *Proceedings of the Data Compression Conference*, pages 140–149, Snowbird, Utha. IEEE Computer Society Press.

Wu, X., Barthel, K. U., and Zhang, W. (1998). Piecewise 2D autoregression for predictive image coding. In *Proceedings of the International Conference on Image Processing (ICIP 1998)*, volume 3, pages 901–904. IEEE Signal Processing Society.

BAYER PATTERN COMPRESSION BY PREDICTION ERRORS VECTOR QUANTIZATION

Antonio Buemi, Arcangelo Bruna, Filippo Vella and Alessandro Capra
Advanced System Technology
STMicroelectronics
Stradale Primosole 50, 95121 Catania
ITALY
Email: (arcangelo.bruna, filippo.vella, antonio.buemi, alessandro.capra)@st.com

Keywords: Bayer Pattern, Color Filter Array (CFA), Vector Quantization (VQ), Differential Pulse Code Modulation (DPCM), Image Compression.

Abstract: Most digital cameras acquire data through a Bayer Colour Filter Array (CFA) placed on sensors where each pixel element records intensity information of only one colour component. The colour image is then produced through a pipeline of image processing algorithms which restores the subsampled components. In the last few years the wide diffusion of Digital Still Cameras (DSC) and mobile imaging devices disposes to develop new coding techniques able to save resources needed to store and to transmit Bayer pattern data. This paper introduces an innovative coding method that allows achieving compression by Vector Quantization (VQ) applied to prediction errors, among adjacent pixel of Bayer Pattern source, computed by a Differential Pulse Code Modulation (DPCM)-like algorithm. The proposed method allows a visually lossless compression of Bayer data and it requires less memory and transmission bandwidth than classic "Bayer-oriented" compression methods.

1 INTRODUCTION

Bayer pattern data (Bayer, 1976) are acquired by camera's sensor used in digital cameras, phone cameras and in many kind of low-cost integrated image acquisition devices. It allows acquiring one colour component for each pixel. The acquired colours are Green, Red and Blue. Then a colour interpolation technique recovers the missed colours. Due to the increasing of the sensor's resolution and to the developing of applications based on the transmission of images, compression of the Bayer pattern has become a strategic feature to reduce the amount of data to store and to transmit.

It's important to point out that any manipulation of the Bayer pattern should preserve the integrity (i.e. the information content) of the data, because every image generation pipeline uses it as input to create the final full colour image. So, ideally, compression of these data should be obtained using lossless techniques, reducing only the inter-pixel and coding redundancies. On the other side, lossless techniques allow low compression rates and preserve psychovisual redundant information.

Lossy compression methods, instead, yield high compression ratios and ideally they should discard only visually not-relevant information.

The problem of the Bayer data compression is quite recent and although traditional coding techniques offer good performances on full colour images, most of them do not offer the same performances with images captured by Colour Filter Array (CFA) digital sensors.

The most trivial, inexpensive solution (both in terms of computational complexity and hardware resources) is to split the Bayer image colour channels and compress them independently using an efficient compression algorithm, i.e. DPCM (Gonzales et alii 1993, Hynix). A more sophisticated compression method for images in Bayer pattern format (Acharya et alii, 2000) is based on Wavelet transform. This approach consists on two steps.

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 283–288.

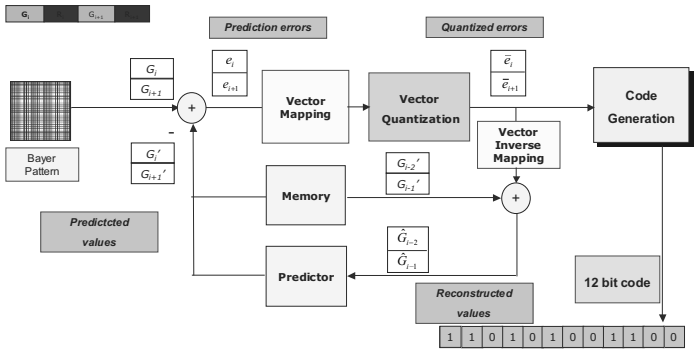


Figure 1: Proposed coding scheme.

First, a 2-dimensional Discrete Wavelet Transform (DWT) is applied and then the DWT coefficients are quantized. DWT is used because it allows to describe abrupt changes better than Fourier transform.

The result is a lossy compression that is perceived as visually lossless by the Human Visual System (HVS). Several sub-band coding compression methods have been introduced by Toi (Toi et alii, 1999) and Le Gall (Le Gall et alii, 1988).

An algorithm (Lee et alii, 2001) for Bayer image compression based on JPEG (Wallace, 1991) has been recently proposed. This method, encodes the image using the JPEG standard. The image is pre-processed to convert from Bayer to YCbCr format with 4:2:2 or 4:2:0 sub-sampling. After this transformation, Y data presents blank pixels, so JPEG compression cannot be directly applied. Therefore a simply 45^{o} rotation of Y data is performed. After the rotation, Y data is localized in the centre of the image in a rhombus shape by removing rows and columns containing blank pixel. The blocks located on the boundaries Y image data are filled by using a mirroring method before encoding with JPEG.

All these approaches require appreciable amount of memory and bandwidth. An efficient compression technique based on Vector Quantization (VQ) techniques coupled with consideration on Human Visual System (HVS) has been proposed recently (Battiato et alii, 2003). Bayer pattern values are gathered in groups of two pixels accordingly to their colour channel and then the generated couple is quantized with a function considering effects of "Edge Masking" and "Luma Masking".

This paper presents an efficient Bayer Pattern lossy compression technique devised specifically to reduce the data acquired by the sensor of 40% with a final low bit rate (about 2:1) and low distortion (about 50 dB of PSNR).

The paper is organized as follows: Section 2 describes the proposed method. Section 3 presents the results of the experiments, while conclusions and final remarks are discussed in Section 4.

2 BAYER PATTERN COMPRESSION

The proposed algorithm scheme is described in Figure 1. First, the DPCM algorithm is applied to compute the difference between the actual and the predicted pixel values gathered in vectors of two consecutive elements of the same colour channel. Prediction is performed assuming the adjacent samples of the same colour components having similar brightness values.

The error between the current value and the prediction is then processed by the "Vector Mapping" block in order to reduce the symmetry and lossy compressed through the "Vector Quantization" block. The vector quantizer has been designed to discard information not perceivable, accordingly to psychovisual considerations based on HVS properties. The quantized values are then taken as input by the "Code Generation" block that will yield a 12 bit code. The encoded data is decoded in order to retrieve the prediction for the next values. This allows avoiding propagation errors during the encoding process.

This scheme assumes to process 10-bpp images, so each couple of 10 bit pixels generates a 12-bit code with a compression of 12/20 (60%). The same process can be easily extended to 8 bpp-images.

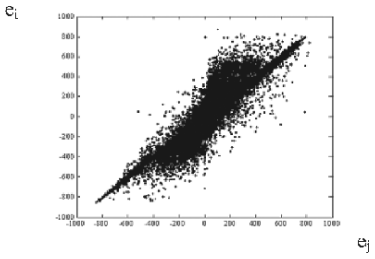
Next subsections describe with more detail each step of the co/decoding procedures.

2.1 Step1:Prediction and Differential Block

The DPCM step is based on the idea of reducing the entropy of the source by coding the difference between the current value and a prediction of the value itself. In our approach the prediction function is performed with 2-dimensional vectors obtained applying VQ. In particular, let (V_i, V_j) be the vector to be coded. The second value of the previous vector (V_{i-1}, V_{j-1}) is used to build a vector of two identical components (V_{j-1}, V_{j-1}) which is the predictor for (V_i, V_j). This strategy has been chosen because V_{j-1} is spatially closer to (V_i,V_j) than V_{i-1} and usually closer samples are statistically more correlated. The error vectors (e_i, e_j) are computed as the difference between the vector to be coded and the prediction vector:

$$(e_i, e_j) = (V_i,V_j)-(V_{j-1},V_{j-1}) = (V_i - V_{j-1}, V_j - V_{j-1}).$$

A typical error distribution for this kind of prediction scheme is showed in Figure 12.



Observe that a very high percentage of values falls near the origin, so the used vector quantizer has been modelled optimizing this distribution.

Note that the prediction function uses, as input, the restored values that are used as input for the predictor block in the decompressor. In this way no errors propagation is present in the compression-decompression system (Figure 1 and Figure 7).

2.2 Step2:Vector Mapping

In the prediction error distribution described above is evident an odd symmetry. It can be exploited to reduce the size of the table used for the VQ. In fact the vectors falling in the third and in the fourth quadrant can be mapped in the first and in the second one, so just the two upper quadrants are to be quantized.

This task is performed by the "Vector Mapping" block: it checks whether the input vector falls in the upper part of the diagram or not. In the first case no changes occur, while, in the other case, the sign of the values is changed as shown in the Figure 3.

One bit is used in the compressed code in order to take into account such mapping (see Paragraph 2.4 for details).

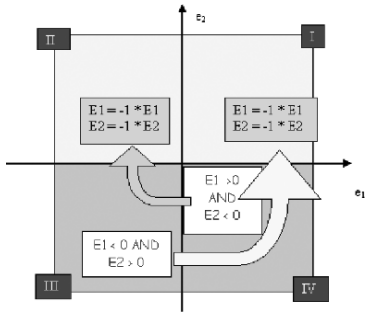


Figure 3: Vector Mapping.

2.3 Step 3:Vector Quantization

Given a vector (X_1,...,X_n) of size N, basic concept of Vector Quantization (Gray et alii, 1998) can be described geometrically. The associated binary representation can be seen as a set of N coordinates locating a unique point in the N-dimensional space. The quantization is performed partitioning the space with N-dimensional cells (e.g. hyperspheres or hypercubes) with no gaps and no overlaps. As the point defined by the input vector falls in one of these cells, the quantization process returns a single vector associated with the selected cell. Finally, such vector is mapped to a unique binary representation, which is the actual output of the vector quantizer. This binary representation (code) can have fixed or variable length.

A vector quantizer is said to be "uniform" if the same quantization step is applied to each vector element, so that the N-dimensional space is divided into regular cells. If the space is partitioned into regions of different size, corresponding to different quantization step, the quantizer is called "not-uniform". The "target" vector is called "codevector" and the set of all codevectors is the "codebook". A grayscale 10 bit image is described by 2-dimensional vectors of brightness values falling into the range [0, 1023]. In the proposed model, each codevector

represents a 2-dimensional input vector. Moreover, the proposed algorithm uses a not uniform quantization according with the properties of HVS. In particular, two considerations should be taken into account: the quantization errors are less visible along the edges and the HVS discriminates better the details at low luminance levels. Thus, in the areas near the origin (where the prediction error is low) a fine quantization is performed and most information is preserved. On the contrary, in the area far from the origin, a coarse quantization is applied due to the presence of boundaries and more information is loss. Furthermore, since DPCM drifts the values towards zero, a very high percentile of input samples will fall in the area around zero.

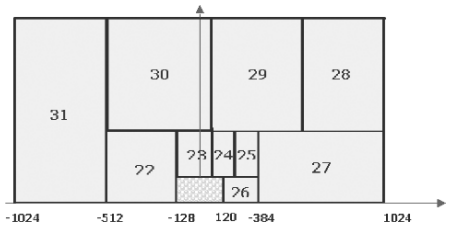


Figure 4: Outer Quantization Regions.

In the testing database used in experiments reported in this paper containing 130 images, the 65% of prediction errors are less than 20 and the 80% of values are less than 38. This information has been exploited partitioning the two upper quadrants of the 2-dimensional space into regions shaped and distributed to minimize the quantization error. Each region has different size and position in the quantization board and it has been divided into 64 "sub-regions". Such regions have been obtained dividing the horizontal and vertical dimension by a constant number. In this way, bigger regions cover bigger areas and the quantization is stronger (more loss of information), while in smaller areas a lighter quantization is applied and most information is preserved (Figure 4). The Figure 5 shows an enlarged image of the quantized region near the origin.

We assumed that the regions are 32 and that each region is fragmented into 64 sub-regions (Figure 6).

Each vector (e_1, e_2) to be quantized is approximated with the nearest couple in the

Figure 5: Inner Quantization Region.

corresponding sub-region. The quantization step in the horizontal direction *X_step* is given by:

$$X\_Step = \frac{RegionWidth}{HorizontalQuantizationStep}$$

In the same way, the quantization step in the vertical direction *Y_step* is given by:

$$Y\_Step = \frac{Region\_Height}{VerticalQuantizationStep}$$

Where *RegionWidth* and *RegionHeight* are the width and the height of each region, while *HorizontalQuantizationStep* and *VerticalQuantizationStep* are the number of partitions in each direction.

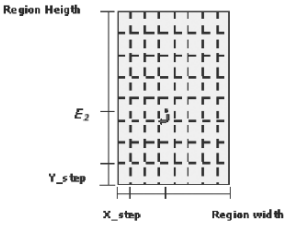


Figure 6: Sub-region partitioning.

The Horizontal and Vertical Quantization Step are set accordingly to the length of the code and the allowed distortion in quantization (Figure 6).

2.4 Step 4: Code Generation

The code representing the vector quantized samples is a fixed length code. It summarizes information

about the vector mapping, the region where the point falls and the quantization steps applied in each region. The first bit is the "Vector Mapping" bit, indicating if the swap between upper and bottom quadrants has happen or not (section 2.2). Next bits following indicate the index of the region in the quantization table. The length of this part of the code depends on the number of regions in the quantization table. The remaining bits give information on the number of steps, in both vertical and horizontal direction inside the region.

In this discussion we assumed that a 12-bits code should be generated, in order to represent samples falling in a space partitioned into 32 regions and 64 "sub-regions". Thus, the code reserves five bits to index 32 regions and 6 bits to index one of the 64 sub-regions. Different code structure could be defined if the space partitioning or the target bitrate change.

2.5 BP Decoding

Decoding procedure consists on three main steps. The first one is the code evaluation allowing the extraction of the compressed values. The second one is the "Inverse Vector Mapping". It assigns the right sign to values depending on the inversion flag. Then the retrieving of the original values is obtained adding the decoded prediction error to the previously restored vector. Since the predictor is equal both in the encoder and the decoder, the predicted values are equal in the two part of the processes, so there is no propagation of error during the decoding process. The decoder block diagram is shown in Figure 7.

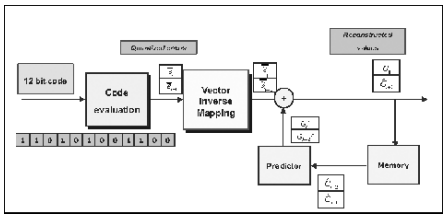


Figure 7: Bayer Pattern Decompressor scheme for a pair of green pels.

3 RESULTS

The algorithm has been tested on two sets of images acquired in different light conditions. A first set, named "Wide Range Image set", contains images acquired by a CMOS-VGA sensor and with a histogram distributed in almost all the intensity range for each channel. In the second set the images have a very narrow histogram. The proposed method had very good performance with a PSNR of about 56 dB in the first set (Figure 8).

Lower, but still high (about 50 dB) PSNR has been achieved when the images with a wider range have been processed (Figure 9). In both cases compression didn't introduce perceptible artefacts in the output images. The algorithm has been compared with another VQ-based compression method introduced by S.Battiato et alii in 2003. Both the approaches yield a fixed-length coding providing a compression from 10 to 6 bpp.

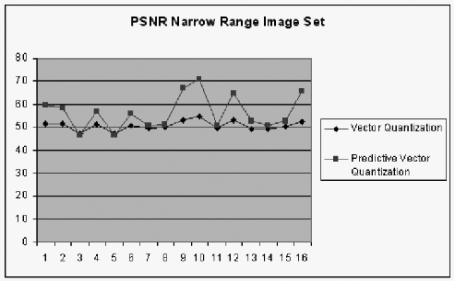


Figure 8: PSNR comparison between the proposed method and a classic algorithm on narrow range image set.

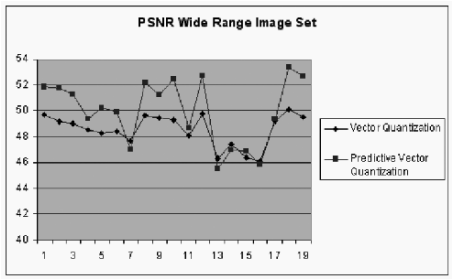


Figure 9: Figure 10: PSNR comparison between the proposed method and a classic algorithm on wide range images.

4 CONCLUSIONS

A new algorithm to compress images acquired by sensors is presented. The algorithm is oriented to the compression of the Bayer pattern, the most used pattern in image acquisition devices.

The technique is based on a predictive schema where the prediction error is encoded. Each input vector is composed by two adjacent pixels of the same colour component.

The proposed method allows achieving a compression of 40% of the input data using a fixed-length 12-bits code assigned to each couple of 10-bits input pixel. Moreover, experimental results showed that compression doesn't involve a perceptible loss of quality in the output. As a result, the bit rate is low and the distortion introduced by compression is very limited (about 50 dB PSNR).

REFERENCES

R.M. Gray, D.L. Neuhoff, 1998, Quantization, *IEEE Trans. on Information Theory*, vol. 44, n. 6.

R.C. Gonzales, R.E. Woods, 1993, Digital Image Processing, Addison Wesley, pp. 358-374.

B.E. Bayer, 1976, Color Imaging Array, U.S. Patent 3,971,065.

Hynix CMOS Image Sensor Application Note 2002, *Endpoints SE401 Demokit Information*, http://www.hynix.com/datasheet/pdf/system_ic/sp/SE401_Demokit.pdf

T. Acharya, L.J. Karam, F. Marino, 2000 Compression of Color Images Based on a 2-Dimensional Discrete Wavelet Transform Yielding a Perceptually Lossless Image, U.S.Patent 6,154,493.

T. Toi, M. Ohita, 1999, A Subband Coding Technique for Image Compression in Single CCD Cameras with Bayer Color Filter Arrays, *IEEE Transaction on Consumer Electronics*, Vol. 45, N. 1, pp. 176-180.

Sang-Yong Lee, A. Ortega, 2001, A Novel Approach of Image Compression in Digital Cameras With a Bayer Color Filter Array, *In Proceedings of ICIP 2001* – International Conference on Image Processing – vol. III 482-485 Thessaloniki, Greece.

G.K. Wallace, 1991, The JPEG still picture compression standard, Communications of the ACM 34(4), pp. 30-44.

Le Gall, A. Tabatabai, 1988, Subband Coding of Digital ImagesUsing Symmetric Shor Kernel Filters and Arithmetic Coding Techniques, in *Proceedings of the ICASSP 88 Conference, (New York)*, pp. 761-764.

S. Battiato, A. Buemi, L. Della Torre, A. Vitali, 2003, "Fast Vector Quantization Engine for CFA Data Compression", *In Proceedings of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing,* NSIP, Grado, Italy.

APPLICATION LEVEL SESSION HAND-OFF MANAGEMENT IN A UBIQUITOUS MULTIMEDIA ENVIRONMENT

Letian Rong and Ian Burnett
CRC SIT (Smart Internet Technology), University of Wollongong, Wollongong, Australia
Email: lr98@uow.edu.au, i.burnett@elec.uow.edu.au

Keywords: Universal multimedia adaptation, Mobile multimedia applications, Internet services and applications.

Abstract: This paper focuses on one of the most important aspects of user mobility in a ubiquitous mobile environment: application session hand-off management. Here we use the term Session Mobility to define the ability of handling application session hand-offs among mobile devices. The paper summarizes the current research in the field and addresses the important facets and the missing "ingredients" of these treatments. We then propose an architecture to support and manage application session transfers based on the MPEG-21 multimedia framework. This takes advantage of Digital Items and adaptation metadata to provide a standards-based approach to the problem. Finally, we validated our framework using a test-bed which provides for dynamic multimedia adaptation.

1 INTRODUCTION

Mobile devices are becoming increasingly popular. In the near future, users will be using a series of devices rather than a single but will also be using them while changing locations rapidly. This is often called personal mobility and is therefore the aim of allowing users to consume services through multiple devices while changing location. This paper focuses on one area of personal mobility, Session Mobility (SM) management, which deals specifically with the issues involved in consuming a session continuously or later, following storage of session state, in a device independent manner.

2 RELATED WORK

To date, the widest published projects in the personal mobility area have been that of the Berkeley ICEBERG project (B. Raman et al., 1999), Stanford's Mobile People Architecture (MPA) (M. Roussopoulos et al., 1999), AT&Ts Telephony Over Packet networkS (TOPS) (N. Anerousis et al., 1999), the Session Initiation Protocol (SIP) (M. Handley et al., 2001), and the Hand-off Manager middleware from University of Illinois and Purdue University (Y. Cui et al., 2003).

The first four of these are similar in that they were designed primarily for voice mail, email and teleconferencing related services. The main feature of this type of application is that they are usually short-lived and require only basic resource adaptations and modality conversions (voice-to-text, voice-to-video with fixed configurations). Thus, the main focus of these systems is to deal with issues involved with user location and they are not designed for the complex adaptation procedures that might be required in multimedia applications.

Cui et al., (Y. Cui et al., 2003) developed a more extensive middleware framework to manage the session hand-off issues of personal mobility missing from the above systems. Importantly, it was also designed to target multimedia applications. Their system supports personal mobility by inserting hand-off manager middleware between multimedia applications and the underlying network infrastructures. They also introduced a User Metadata Server for storing session state related information. Their hand-off manager is then responsible for transferring the session state of the primary application to the hand-off manager of another application directly or alternatively to instruct the User Metadata Server to perform the operation. They also introduced a QoS mapping algorithm to facilitate resource transcoding during session hand-offs.

There are, however, several limitations to Cui et al's contribution (Y. Cui et al., 2003). First the transcoding-based QoS adaptation approach from the hand-off manager middleware is too complex and is limited to one specific element of the multimedia

J. Ascenso et al. (eds.), e-Business and Telecommunication Networks, 289–295.

content adaptation process (i.e., transcoding). A more comprehensive adaptation process is therefore required to support dynamic adaptation of multimedia resources and sessions. This is particularly important during the process of session hand-off amongst a broad variety of devices, and the middleware should be able to support universal multimedia in particular, user preference, terminal capability, network capability, natural environment characteristics and other requirements.

In addition, the Cui approach takes a session sender approach that requires a session to be stopped before the hand-off of the session can occur. In certain cases the consumer might like to have data seamlessly streamed from one device to another by initiating it from the session receiver. On other occasions the consumer might like to continue the same session on multiple devices. For example, a person might like to continue watching a movie on his PDA while his kids continue watching the same movie on the TV. Cui's approach does not facilitate these possibilities.

Finally, all reported approaches to session mobility so far deal only with one to one session hand-off. In real life situations, however, it is quite possible that the hand-off will involve N to M session transfers. For example, the session state of a presentation may need to be synchronized to other devices so that others can either carry on the presentation or view the updated presentation session. It is also important that a device, on receiving several sessions from multiple session senders try to conserve their previous application session states (i.e., window size, window position etc.,). This requires a more complex session hand-off manager.

The architecture proposed in this paper addresses these issues and facilitates session hand-off with particular focus on the following requirements:

1. Enabling sessions to be transferred amongst a large variety of devices and allowing those sessions to be adapted to the new environment dynamically. This involves resource(s) adaptation according to the specific usage environment and session states. Furthermore, a mechanism is required to identify which application data are session related (e.g., how long a movie has been playing) and which are usage environment related (e.g., the resolution of a movie). This would enable the device to identify the application data which should be transferred to the session receiver.
2. Enabling sessions to be stored remotely for latter consumption or to be transferred seamlessly between devices. The management software should intelligently support those options.
3. Enabling N to M session transfers and managing multiple received sessions in an appropriate way so as to conserve their previous session states.

3 SESSION MOBILITY ARCHITECTURE

The proposed session mobility architecture (see Figure 1) was designed based on the MPEG-21 multimedia framework. This approach takes advantage of Digital Items and adaptation metadata to provide a standards-based approach to the problem. The architecture consists primarily of two parts: the M21 middleware which is installed on all the consumer devices and the SM server. We shall now discuss MPEG-21 and the roles of these two parts of the architecture.

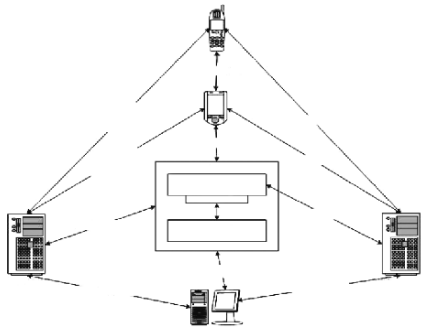


Figure 1: Session mobility architecture.

3.1 MPEG-21

MPEG-21 is a new multimedia framework from the Moving Picture Experts Group (MPEG) that supports multimedia access and delivery using heterogeneous networks and terminals in an inter-operable and highly automated manner (J. Bormans et al., 2003). MPEG-21 addresses the requirements of Universal Multimedia Access (UMA) (A. Perkis et al., 2001) by providing a normative open framework for multimedia delivery and consumption.

The fundamental unit of distribution and transaction in the MPEG-21 framework is the Digital Item (DI). It can be considered as a structured digital object which consists of resource(s) (e.g., a photo album, a web page) and related information for the manipulation of the resource(s) (e.g., terminal capabilities, intellectual properties). DI is represented as a Digital Item Declaration (DID) in MPEG-21 through the Digital

Item Declaration Language (DIDL) (V. Iverson et al., 2002) which conforms to the XML standard. While a DI can have wide ranging contents (e.g. Rights information), for our purposes we consider a DI to contain resource(s), i.e. a list of choices that correspond to the various adaptation aspects of those resources and Digital Item Adaptation (DIA) information. DID (ISO 21000-2) has already become an International Standard.

The MPEG-21 Digital Item Adaptation (DIA) Tools are a collection of descriptions and format independent mechanisms that steer the multimedia content adaptation process (A. Vetro et al., 2004). The descriptors are represented in XML and can be either wrapped in a DID or be used independently. Currently the DIA Tools are clustered into eight major categories (see Figure 2) and our architecture mainly uses three of those tools: Usage Environment Description Tools, Session Mobility and DIA Configuration Tools.

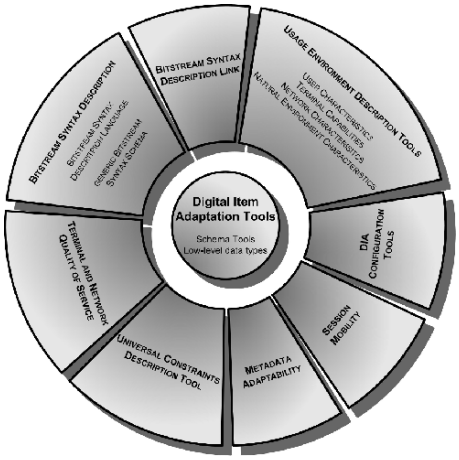


Figure 2: DIA tools architecture (A. Vetro et al., 2004).

The Usage Environment Description Tools includes descriptors for the various dimensions of the usage environment. Those are user characteristics, terminal capabilities, network characteristics and natural environment characteristics.

The DIA Configuration Tools specifies how and where the related usage environment information can be used for the adaptation of DIs, also it identifies which configuration state of resources are session related or should be configured again during session hand-offs.

Session Mobility describes the application session state information that pertains to the consumption of a Digital Item in real time, allowing a Digital Item to be consumed continuously when it is transferred from one device to another.

The usage of these related tools are explained in the following section. Since the scope of MPEG-21 is very broad, interested readers are referred to references for information related to MPEG-21 (J. Bormans et al., 2003) (A. Perkis et al., 2001).

3.2 M21 middleware

M21 is a middleware application developed at the University of Wollongong that implements the base concepts of MPEG-21. It was used in our previous work (L. Rong and I. Burnett, 2004) for enabling multimedia contents to be adapted to various devices/terminals in a ubiquitous environment according to usage environment attributes. Its structure is illustrated in Figure 3.

In the multimedia content adaptation approach, the DI was used as a "Menu" to contain a link or links to multiple pre-existing variations of resources and a list of choices which correspond to the resource or which can be selected so as to configure those resources further. Furthermore, the DIAC related information is also encapsulated in the DI to provide further guidelines on the adaptation process. The guidelines specify information on the location of adaptations e.g. at the consumer device, intermediate nodes or the provider, the types of descriptors required for adaptations and how the choices should be selected (by the device automatically or by the user manually). The consumer then requests a multimedia resource by first requesting its corresponding Digital Item and then performs the resource related configuration and selection according to his/her usage environment attributes and the "guidelines" in the DI. The usage environment attributes are expressed in the Usage Environment Description Tools. The consumer then sends that configuration and selection information as the second request to the provider who in turn, performs the required resource related adaptations on their side (according to the same DI) before sending the resource to the consumer. The M21 middleware is installed on both the consumer and provider side to perform the DID parsing, DIA processing and resource adaptation processes. Also it handles launching of applications to "play" the resources on the consumer side.

We adopted the existing M21 middleware and the content adaptation approach into the session mobility architecture. This gives us two main advantages. The

first advantage is that we can use the existing content adaptation functionality to reconfigure and adapt resources as they are transferred to a session receiver device. The second advantage is that the Digital Item is used as the fundamental unit for multimedia delivery in the session state transfer process. This broadens the concept of session state because a DI can then consist of multiple resources and hence the session state related to a DI can contain multiple media session states – each associated with media resources in the DI. Thus, we define the session state of a DI as a user's current state of interaction with a DI. An example of this concept would be the session state of a digital CD album which contains audio clips, movies and photos (i.e., a DI with multiple resources). The session state of the DI (in our case the CD) would contain the audio track playing position, the movie playing position, the photo being displayed etc. The session state of a DI is inserted into the DI as a SM schema of the DID Adaptation Tools when it is transferred to the session receiver device. Also DIAC is used in the DI to identify which configuration state is session related.

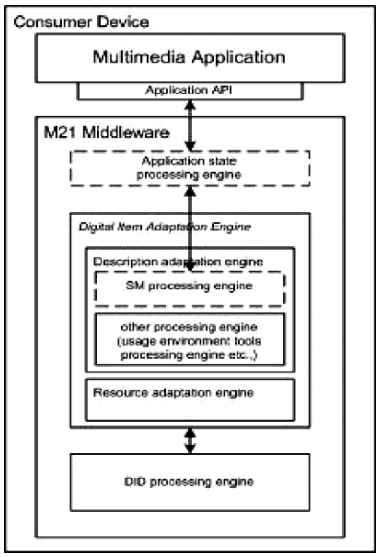


Figure 3: M21 middleware and its components.

A SM processing engine and application state processing engine are thus added into the middleware to manage session hand-offs (see as the dot line boxes in Figure 3). The SM processing engine is used for generating and parsing session state information as SM schemas, and it also acts as the decision maker as to where sessions should be fetched from. Its decision making mechanism is explained in section 3.4. It should be noted here that the SM processing engine links directly with the application state processing engine by fetching application session information from it on the session sender side and passing this information to the engine on the session receiver side.

The application state processing engine is used to extract application session information through the application APIs on the session sender side and launching appropriate applications on the receiver side. As mentioned above, the engine connects to the SM processing engine to transfer the information, so that session states can be encapsulated/extracted from DIs as SM schemas. Again, we believe it is necessary to support multiple session transfers. The application state processing engine performs this task by launching appropriate applications with resources in their preserved session state and arranging those applications to reflect their previous application session states. It therefore takes advantage of the middleware architecture and only requires different application "drivers" to support various applications through their APIs.

3.3 SM Server

The second major part of our architecture is the Session Mobility (SM) server that resides in the network for storing, redirecting and processing multimedia sessions. It uses a database to store session related information as indicated in Table 1.

The SM server either transfers stored sessions directly to any session receivers or directs session senders to perform the operation. Its session steering mechanism is explained in the section below.

Table 1: Session Mobility server information.

User ID	User identification
IP Address	IP address of the device
Application Name	Application for running the media
Resource Name	Name of media resource
Media Status	Playing, Paused or Stopped
Stored Session	Exist or empty

3.4 Multimedia session hand-off management

We categorize session hand-offs into two main types, session sender driven and session receiver driven. The two approaches are shown in Figures 4 and 5.

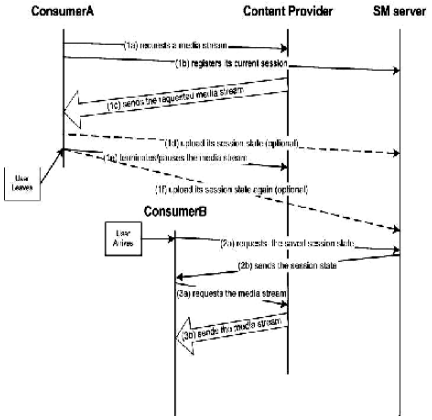


Figure 4: Session sender driven approach.

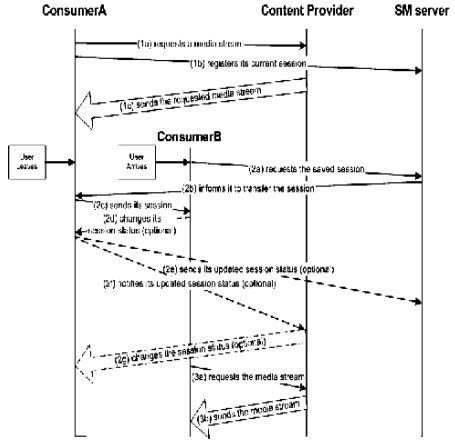


Figure 5: Session receiver driven approach.

Sender driven session hand-off is when the session sender steers the session hand-off process. In this case the sending terminal changes its current session status (i.e., play, pause or stop) before transferring its session to the session receiver or the SM server. The alternative, receiver driven approach, enables the session receiver to control the session status of the session sender and initiate the session hand-off process. The steps of each algorithm are shown in Figure 4 and 5.

In the steps detailed, a content provider is used in the case of streaming media sessions being transferred during hand-offs. Those steps however can equally be used for transferring local data without the content provider.

Both approaches have been incorporated into our current architecture and a simple algorithm is used by the SM processing engine for determining which approach it should use according to the current status of the media. If a media resource is currently playing and there are no corresponding sessions saved on the SM server, then the session receiver would contact the session sender to retrieve the media session directly. Conversely if the media session has already been saved on the SM server and it shows the media status as being stopped, paused or killed, then the session receiver would retrieve the session from the SM server instead. The advantages of facilitating both approaches in the architecture are that: 1. Consumers are given the flexibility to initiate session hand-offs from different devices; 2. It reduces SM server load by offsetting some session hand-offs to peer-to-peer session transfers (i.e., in the session receiver driven approach). Finally, several enhanced approaches from (Y. Cui et al., 2003) can be easily modified to be incorporated into our proposals.

4 EXPERIMENTAL SETTINGS AND VALIDATION

The M21 middleware was installed on a test-bed as the middleware layer application on three separate computers and a fourth was used as both the SM server and RTP streaming server.

Two different applications were used in the experiment to demonstrate multiple session management. These were a RTP streaming client and a Java web browser. The RTP streaming program pair, a RTPTransmitter and RTPReceiver, have been used previously in our dynamic multimedia adaptation work (L. Rong and I. Burnett, 2004), and the Java web browser is implemented based on NetClue (NetClue Java web browser v4.2) APIs. All programs were written in Java for common application interaction and easy code deployment.

Both session sender driven and session receiver driven approaches with single or multiple session transfer(s) have been evaluated in the test-bed. The session sender devices were able to register their current sessions with the SM server and transfer their session to either the SM server or session receiver devices through the M21 middleware. The session receiver devices were then able to launch corresponding applications, adapt contents according to their usage environment attributes and resume processing of the Digital Item session. The session receivers can also effectively control the media status of the session

sender through the M21 middleware as shown in Figure 6.

Furthermore the application session processing engine arranges and displays applications according to their previous application states. Currently the supported application session information includes window size, position, state i.e., maximized, minimized or user defined and size-to-resolution ratio.

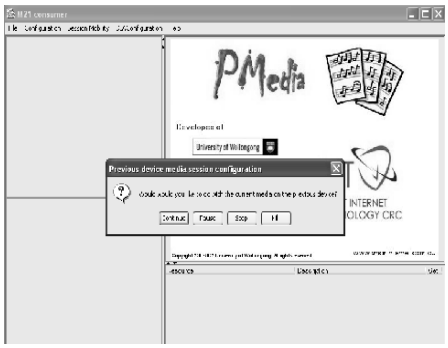


Figure 6: screen shot shows session receiver choosing media status of session sender through M21.

In addition, the test-bed supports the following usage environment descriptors and choices for streaming medias: MediaTime, Coarse Language, CharacterSet and Resolution. The descriptors MediaTime and Coarse Language are transferred to the session receiver device as they are session related, while CharacterSet and Resolution are configured locally according to the capabilities of the session receiver. For the web browser, content adaptation was demonstrated by adapting web pages of different resolutions to various terminals according to their resolutions. The generation and processing overheads of MPEG-21 Digital Items are shown to be relatively low in the experiment. We used 83ms and 2ms on average for generating Digital Items with application session state information for the RTP streaming client and web browser respectively, and 279ms for parsing the DIs on the receiver end. The generation of DI for the RTP streaming client is relatively higher due to fact that session information is required to be extracted from the client program during media streaming.

5 CONCLUSION

In this paper, we proposed a session mobility architecture to improve the area of session hand-off management with the aim of targeting multimedia applications. There are some similarities between our work and that of Cui's. We believe that we have substantiated what is missing from their work and broader session transfer experience with MPEG-21.

This architecture enables session transfers to be performed through two different types of approaches (session sender driven and session receiver driven) and this then facilitates different types of session transfer needs. We also adopt the Digital Item concept to facilitate dynamic session adaptations to the session receiver, without complex content negotiation/matching algorithms. Further, an application state processing engine with different "application drivers" is used to manage multiple applications during session transfers. This can be easily expanded to cater for other multimedia applications by writing application "drivers" and incorporating them into the application state processing engine.

Several Digital Item Adaptation Tools (i.e., DIA Configuration Tools, Session Mobility and some of Usage Environment Description Tools), are the end-result of our work and they are now included in the Final Draft International Standard (FDIS) version of DIA Tools and will become an International Standard in 2004.

As for our future work, we shall implement a more complex Session Mobility engine that uses more comprehensive mobility characteristic descriptors (Z. Sahinoglu and A. Vectro et al., 2003) for facilitating session hand-offs and their adaptation processes under more complex mobility situations.

REFERENCES

B. Raman, R. H. Katz and A. D. Joseph, "Personal Mobility in the ICEBERG Integrated Communication Network", Report No. UCB/CSD-99-1048, May, 1999.

M. Roussopoulos, P. Maniatis, E. Swierk, K. Lai, G. Appenzeller and M. Baker, "Person-level Routing in the Mobile People Architecture", 2nd USENIX Symposium on Internet Technologies and Systems (USITS 1999), October, 1999.

N. Anerousis, R. Gopalakrishnan, C. R. Kalmanek, A. E. Kaplan, W. T. Marshall, P. P. Mishra, P. Z. Onufryk, K. K. Ramakrishnan, and C. J. Sreenan, "TOPS: An Architecture for Telephony Over Packet Networks", IEEE Journal of Selected Areas in Communications, 17(1), January 1999.

M. Handley, H. Schulzrinne, E. Schooler and J. Rosenberg, "SIP: Session Initiation Protocol", RFC2543, April, 2001.

Y. Cui, K. Nahrstedt and D. Xu, "Seamless User-level Hand-off in Ubiquitous Multimedia Service", Multimedia Tools and Applications Journal, Special Issue on Mobile Multimedia and e-Commerce, Kluwer, 2003.

J. Bormans, J. Gelissen, and A. Perkis, "MPEG-21: The 21st Century Multimedia Framework", IEEE Signal Processing Magazine, March 2003.

A. Perkis, Y. Abdelijaoued, C. Christopoulos, T. Ebrahimi, "Universal Multimedia Access from Wired and Wireless systems", Birkhauser Boston, transactions on Circuits, Systems and Signal Processing, Special issue on Multimedia Communications, vol. 20., No. 3, 2001, pp. 387-402.

V. Iverson, T. Schwartz, Y. Song, R. Walle, E. Santos and D. Chang, "MPEG-21 Digital Item Declaration FDIS", ISO/IEC JTC1/SC29/WG11/N4813, May, 2002.

A. Vetro, S. Devillers and C. Timmerer, "ISO/IEC 21000-7 FDIS – Part 7: Digital Item Adaptation", ISO/IEC JTC1/SC29/WG11/N6168, January 2004.

L. Rong and I. Burnett, "Dynamic Multimedia Adaptation with MPEG-21", Consumer Communication and Networking Conference (CCNC 2004), January, 2004.

NetClue Java web browser v4.2, NetClue, www.netcluesoft.com.

Z. Sahinoglu and A. Vectro, "Mobility Characteristics for Multimedia Service Adaptation", EURASIP Journal of Image Communications, Special Issue on Multimedia Adaptation, October, 2003.

AUTHOR INDEX