Lecture Notes in Electrical Engineering 302

Lieven De Strycker
Editor

ECUMICT 2014

Proceedings of the European Conference on the Use of Modern Information and Communication Technologies, Gent, March 2014

Springer

Lecture Notes in Electrical Engineering

Volume 302

Board of Series Editors

Leopoldo Angrisani, Napoli, Italy
Marco Arteaga, Coyoacán, México
Samarjit Chakraborty, München, Germany
Jiming Chen, Hangzhou, P.R. China
Tan Kay Chen, Singapore, Singapore
Rüdiger Dillmann, Karlsruhe, Germany
Gianluigi Ferrari, Parma, Italy
Manuel Ferre, Madrid, Spain
Sandra Hirche, München, Germany
Faryar Jabbari, Irvine, USA
Janusz Kacprzyk, Warsaw, Poland
Alaa Khamis, New Cairo City, Egypt
Torsten Kroeger, Stanford, USA
Tan Cher Ming, Singapore, Singapore
Wolfgang Minker, Ulm, Germany
Pradeep Misra, Dayton, USA
Sebastian Möller, Berlin, Germany
Subhas Mukhopadyay, Palmerston, New Zealand
Cun-Zheng Ning, Tempe, USA
Toyoaki Nishida, Sakyo-ku, Japan
Federica Pascucci, Roma, Italy
Tariq Samad, Minneapolis, USA
Gan Woon Seng, Nanyang Avenue, Singapore
Germano Veiga, Porto, Portugal
Junjie James Zhang, Charlotte, USA

For further volumes:
http://www.springer.com/series/7818

About this Series

"Lecture Notes in Electrical Engineering (LNEE)" is a book series which reports the latest research and developments in Electrical Engineering, namely:

- Communication, Networks, and Information Theory
- Computer Engineering
- Signal, Image, Speech and Information Processing
- Circuits and Systems
- Bioengineering

LNEE publishes authored monographs and contributed volumes which present cutting edge research information as well as new perspectives on classical fields, while maintaining Springer's high standards of academic excellence. Also considered for publication are lecture materials, proceedings, and other related materials of exceptionally high quality and interest. The subject matter should be original and timely, reporting the latest research and developments in all areas of electrical engineering.

The audience for the books in LNEE consists of advanced level students, researchers, and industry professionals working at the forefront of their fields. Much like Springer's other Lecture Notes series, LNEE will be distributed through Springer's print and electronic publishing channels.

Lieven De Strycker
Editor

ECUMICT 2014

Proceedings of the European Conference on the Use of Modern Information and Communication Technologies, Gent, March 2014

Springer

Editor
Lieven De Strycker
Faculty of Engineering Technology
KU Leuven
Gent
Belgium

ISSN 1876-1100 ISSN 1876-1119 (electronic)
ISBN 978-3-319-05439-1 ISBN 978-3-319-05440-7 (eBook)
DOI 10.1007/978-3-319-05440-7
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014933214

© Springer International Publishing Switzerland 2014
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The ECUMICT conference is organized in order to offer a forum to researchers, where they can present recent results of applied research preferably with a strong link to practical implementations or to industrial applications and recent research results from cooperation between industrial and academic partners.

For this sixth edition we have chosen to focus again on Wireless and Mobile Applications, which is the broad research topic of the organizing research group DraMCo. Mobile and wireless devices are ubiquitous: eye-catching powerful devices offering increasingly sophisticated services, hidden lightweight nodes handling more and more tasks in the background and embedded devices giving wireless communication and network possibilities to a rising number of objects.

These proceedings collect the papers that have been accepted for presentation on the sixth ECUMICT 2014 conference.

I wish to thank everybody who helped us to turn this conference into a success. First of all, I would like to thank the keynote speakers who will present one of their current research topics at the beginning of each session: drWout Joseph, dr Ivan Ndip and dr Claude Oestges. The members of the scientific committee have proven their expertise in announcing and recommending our conference, in reviewing the papers and in presenting a lot of valuable suggestions to the authors. Many colleagues at our Campus and our international partner institutions helped to prepare and distribute the call for papers, invested a significant amount of time in composing these proceedings, and were involved in the practical organization of the conference. I also wish to thank the colleagues from our Flemish partner institutes for their enthusiastic help in organizing the second Engineering Technology Electronics-ICT doctoral symposium. And, last but not least, I want to thank the authors: only with their contributions, this sixth ECUMICT conference could be successful.

Gent
3rd February 2014

Editor
Lieven De Strycker
KU Leuven @ KAHO Sint-Lieven
Faculty of Engineering Technology
Gebroeders de Smetstraat 1, 9000 Gent, Belgium

Contents

Optimization of the POWER Pulse 1
Nicolae Dumitru Alexandru, Alexandra Ligia Balan

Influence of Impulse Noise on Alamouti Code Performances 11
Mihaela Andrei, Lucian Trifina, Daniela Tarniceriu

Making OpenID Mobile and Privacy-Friendly 23
Faysal Boukayoua, Karel Dewitte, Vincent Naessens

ErasmusApp: A Location-Based Collaborative System for Erasmus Students 35
Karel Bruyneel, Benedita Malheiro

Smart Object for 3D Interaction 49
Hannes Harms, Toomas Juht, Anna Janaszkiewicz, Jana Valauskaité, António Silva, Benedita Malheiro, Cristina Ribeiro, Manuel Silva, Nídia Caetano, Paulo Ferreira, Pedro Guedes

Study of a Wake Up Radio Architecture for Home Multimedia Networks 63
Aissa Khoumeri, Florin Hutu, Guillaume Villemaud, Jean-Marie Gorce

Hybrid Multi-objective Network Planning Optimization Algorithm 73
Ning Liu, David Plets, Wout Joseph, Luc Martens

A Context-Aware Framework for Media Recommendation on Smartphones 87
Abayomi M. Otebolaku, Maria T. Andrade

Measuring the NFC Peer-to-Peer Data Rate 109
Geoffrey Ottoy, Sam Van Den Berge, Jean-Pierre Goemaere, Lieven De Strycker

A Physical Model for Predicting throughput of Wireless LANs 123
Mostafa Pakparvar, David Plets, Luc Martens, Wout Joseph

Study of New Organic Field Transistors for RFID, Optoelectronic and Mobile Applications 135
Marius Prelipceanu, Adrian Graur

Applicability of Amdahl's Law in Multisession TCP/IP Communication 143
Radu-Cezar Tarabuta, Alin Potorac, Doru Balan, Adrian Graur

Fine-Tunning a MAP Error Correction Algorithm for Five-Key Chording Keyboards 153
Adrian Tarniceriu, Bixio Rimoldi, Pierre Dillenbourg

Design of a Modular Simulation Environment for Vehicle Mounted Logical Units 167
Christoph Uran, Helmut Wöllik

Out-of-Band Password Based Authentication towards Web Services 181
Jan Vossaert, Jorn Lapon, Vincent Naessens

Measurement Based Indoor Radio Channel Modeling and Development of a Fading Optimized Circular Polarized Patch Antenna for Smart Home Systems within the SRD Band at 868 MHz . . . 193
S. Wunderlich, M. Welpot, I. Gaspard

Influence of Bluetooth Low Energy on WIFI Communications and Vice Versa 205
Jeroen Wyffels, Jean-Pierre Goemaere, Bart Nauwelaers, Lieven De Strycker

Gestural Interfaces for Mobile and Ubiquitous Applications 217
Ionut-Alexandru Zaiti, Stefan-Gheorghe Pentiuc

Author Index 231

Optimization of the POWER Pulse

Nicolae Dumitru Alexandru and Alexandra Ligia Balan

Abstract. An optimization of the performant inter-symbol interference (ISI) free POWER pulse generated by an improved Nyquist filter with a power-exponential frequency characteristic is presented. The performance of the POWER pulse is studied with respect to the ISI error probability, in the presence of timing error. In this article we show the necessity to apply an optimization scheme for the presented results, and provide an analysis of the optimized results. Furthermore, we identify some challenge in implementing the studied POWER pulse.

1 Introduction

Over the past few years a significant number of studies on better bandwidth reuse and higher error-free data rates for digital communications was published. Improved Nyquist filters are primary used to provide inter-symbol interference free pulses. The raised cosine (RC) pulse is probably the most popular Nyquist pulse because of the transmission without distortion the presence of inter-symbol interference (ISI). Several recent works [3] – [9] dealt with improved Nyquist filters (INFs) that have been successfully used in order to achieve a more open receiver eye with smaller maximum distortion and a smaller probability of error in the presence of timing jitter than the RC pulse for the same excess bandwidth.

The pioneering works on this subject are due to Franks [1] and [2] (1977). The improved Nyquist pulse was introduced by Beaulieau et al. [3] and further dealt with in [4]. Several important contributions were brought in [5]- [14].

More recently an ISI-free power pulse with one parameter for a given excess bandwidth was proposed by Mohri and Hamamura [13]. It has been shown that the POWER pulse performs better in terms of error probability in the most common cases. However his design parameter was not optimized to minimize the error probability for all the considered cases. We investigated it to find the optimal parameter for a wider range of scenarios in order to identify its importance for practical use.

Nicolae Dumitru Alexandru
"Gheorghe Asachi" Technical University of Iaşi, 700506, Romania

Alexandra Ligia Balan
Stefan cel Mare University of Suceava, 720229, Romania
e-mail: alexandra@eed.usv.ro

L. De Strycker (ed.), *ECUMICT 2014*,
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_1, © Springer International Publishing Switzerland 2014

2 Design Aspects Concerning the POWER Pulse

In the following section, we overview the POWER pulse, highlighting advantages and disadvantages of its use. The POWER pulse was proposed an investigated in [13]. It is defined in the frequency domain as

$$S_{POWER}(F)=\begin{cases}1, & |f|\le B(1-\alpha)\\ 1-\frac{1}{2}\left[\frac{|f|-(1-\alpha)B}{\alpha B}\right]^{\beta}, & B(1-\alpha)\le|f|<B\\ \frac{1}{2}\left[\frac{(1+\alpha)B-|f|}{\alpha B}\right]^{\beta}, & B<|f|\le B(1+\alpha)\\ 0, & B(1+\alpha)<|f|\end{cases} \tag{1}$$

Here B is the cut-off frequency, α, ($0\le\alpha\le1$), is the excess bandwidth, and β, ($\beta\ge0$), is the roll-off tuning factor that shapes the POWER pulse.

The frequency characteristic of the POWER pulse is represented in Fig.1 for several values of β.

One can infer from Fig.1 that with increasing values of β less energy is transferred within the transition region from the area near $B(1-\alpha)$ to the area below $B(1+\alpha)$.

Also analyzing equation 1 it was seen that for small integer values of β the pulse shape defined in frequency is simplified, therefore the resulting waveform in the time domain has a known structure.

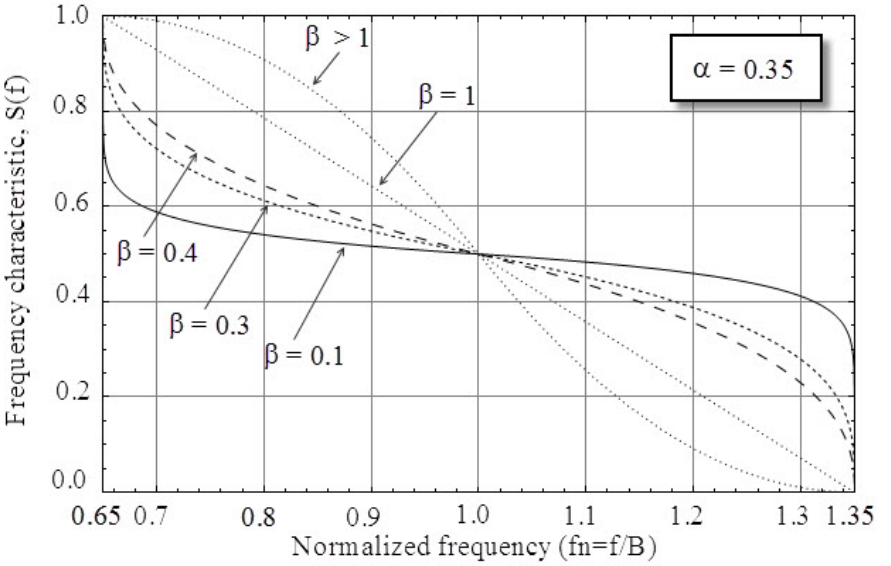


Fig. 1 Frequency characteristic of the POWER pulse for

The POWER impulse response [13] is obtained as

$$s_p(t)=\frac{\sin(2\pi Bt)}{\pi t}\left\{1-\frac{4\pi^2\alpha^2B^2t^2}{2+3\beta+\beta^2}\,{}_1F_2\left(1;\frac{3+\beta}{2},\frac{4+\beta}{2};-\pi^2\alpha^2B^2t\right)\right\} \quad (2)$$

where $B=1/2T$, α, ($0\le\alpha\le1$), is the excess bandwidth, β, ($\beta\ge0$), is the roll-off tuning factor that shapes the POWER pulse in order to get best performance in terms of error probability when sampled with a timing error, and ${}_1F_2(x;a,b;y)$ is the hypergeometric function defined as

$${}_1F_2(x;a,b;y)=\sum_{k=0}^{\infty}\frac{(x)_k}{(a)_k(b)_k}\cdot\frac{y^k}{k!} \quad (3)$$

where $(u)_0=1$, and $(u)_k=u(u+1)(u+2)...(u+k-1)$.

Note that due to the infinite sum in (3) this must be truncated for practical reasons. Unfortunately, there are difficulties in the implementation of the ${}_1F_2(x;a,b;y)$ hypergeometric function.

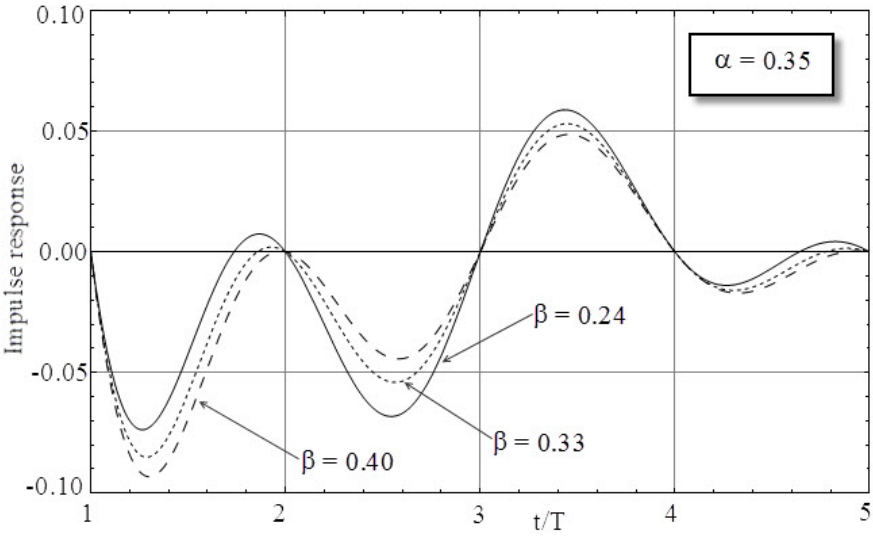


Fig. 2 Time characteristics of the POWER pulse

Figure 2 illustrates the impulse response of the POWER filter for $\alpha = 0.35$ and $\beta = 0.24, 0.33,$ and 0.40 .The values of β are the optimized ones for timing offsets of 0.05, 0.1, and 0.2.

The main motivation of the following analysis is that a pulse with the minimum number of design parameters and smallest BER for any arbitrary value of the roll-off factor and sampling time error, represents the key goal in the design of practical receivers. [9]

3 Optimization of the POWER Pulse

Due to the difficulties in the implementation of the ${}_1F_2(x;a,b;y)$ hypergeometric function the optimization performed here was based on varying the design variable β in a limited region for values of α that takes discrete values in the interval $(0,1]$ with a step size of 0.01, which is convenient for practical purposes. We selected the value of β that results in minimal value of the error probability when the impulse response is sampled with a timing offset.

The error probability P_e was determined using Fourier series as in [15]:

$$P_e = \frac{1}{2} - \frac{2}{\pi} \sum_{\substack{m=1 \\ M\ odd}}^{M} \left(\frac{\exp\left(-m^2\omega^2/2\right)\sin\left(m\omega g_0\right)}{m} \right) \prod_{k=N_1}^{N_2} \cos\left(m\omega g_k\right) \qquad (4)$$

Here M represents the number of coefficients considered in the approximate Fourier series of noise complementary distribution function; $\omega = 2\pi/T_f$ - angular frequency; T_f is the period used in the series; N_1 and N_2 are the number of interfering symbols before and after the transmitted symbol; $g_k = p(t - kT)$ where $p(t)$ is the impulse response used for transmission, and T is the bit duration. The results are computed using $T_f = 40$ and $M = 61$ for $N=2^8$ interfering symbols and SNR=15 dB. Usually the calculations are performed for $N=2^9$ interfering symbols, but we used $N=2^8$ in order to avoid major difficulties due to the implementation of the ${}_1F_2(x;a,b;y)$ hypergeometric function in computing the error probability.

Table 1 summarizes the minimum values found for ISI error probability, when α takes values in the interval $(0,1]$ with a step size of 0.01, which is convenient for practical purposes. We selected the value of β that produces a minimal value of the error probability when the impulse response is sampled with timing offset.

Therefore, our main goal is to optimize the values of the design parameters which interfere in the filter definition, in order to obtain minimal values of the error probability.

Furthermore, we apply an optimization scheme and derived a cubic relationship that exists between the shape tuning parameter β, and the excess bandwidth factor α, for fixed values of t/T.

The optimization problem can be efficiently resolved by implementing some of the well-known numerical optimization methods.

The optimization of the pulse was performed using the numerical optimization Nelder-Mead method that is computationally compact.

Table 1 Optimized values of design parameter β

α	t/T=0.05		t/T=0.1		t/T=0.2	
	β_{opt}	Pe	β_{opt}	Pe	β_{opt}	Pe
0.05	0.02	1.42377*10^-7	0.04	9.46801*10^-6	0.03	3.40438*10^-3
0.10	0.07	1.00761*10^-7	0.10	4.56017*10^-6	0.10	1.68129*10^-3
0.15	0.11	7.4614*10^-8	0.16	2.39533*10^-6	0.17	8.09439*10^-4
0.20	0.15	5.73361*10^-8	0.21	1.35896*10^-6	0.23	3.97072*10^-4
0.25	**0.18**	**4.54259*10^-8**	**0.25**	**8.24052*10^-7**	**0.29**	**2.03023*10^-4**
0.30	0.21	3.69125*10^-8	0.29	5.29313*10^-7	0.34	1.09361*10^-4
0.35	**0.24**	**3.06277*10^-8**	**0.33**	**3.56401*10^-7**	**0.40**	**6.22455*10^-5**
0.40	0.27	2.58633*10^-8	0.35	2.47834*10^-7	0.44	3.73118*10^-5
0.45	0.30	2.21984*10^-8	0.36	1.75243*10^-7	0.48	2.32797*10^-5
0.50	**0.33**	**1.93824*10^-8**	**0.37**	**1.25042*10^-7**	**0.52**	**1.49477*10^-5**
0.55	0.37	1.72671*10^-8	0.40	9.12048*10^-8	0.54	9.85244*10^-6
0.56	0.38	1.69158*10^-8	0.40	8.60203*10^-8	0.55	9.10698*10^-6
0.57	0.39	1.65854*10^-8	0.41	8.12639*10^-8	0.55	8.43271*10^-6
0.60	0.42	1.57092*10^-8	0.44	6.92589*10^-8	0.58	6.76875*10^-6
0.65	0.48	1.45703*10^-8	0.49	5.49627*10^-8	0.63	4.87614*10^-6
0.70	0.55	1.37455*10^-8	0.55	4.54733*10^-8	0.68	3.67259*10^-6
0.75	0.62	1.31589*10^-8	0.62	3.91166*10^-8	0.73	2.87357*10^-6
0.80	0.69	1.27558*10^-8	0.69	3.48871*10^-8	0.77	2.32296*10^-6
0.85	0.78	1.24939*10^-8	0.76	3.21568*10^-8	0.81	1.93373*10^-6
0.90	0.86	1.23416*10^-8	0.84	3.05106*10^-8	0.85	1.65771*10^-6
0.95	0.95	1.22742*10^-8	0.93	2.96595*10^-8	0.90	1.46756*10^-6
1.00	**1.03**	**1.22723*10^-8**	**1.02**	**2.94085*10^-8**	**0.96**	**1.34535*10^-6**

Our goal was to minimize the error probability for fixed values of the timing offset, and determine the optimal values of the parameter β as a function of the excess bandwidth α.

We used β as free parameter (FP). The optimal value of shape tuning parameter β was found to be in cubic relationship with the excess bandwidth α.

$$\beta(\alpha) = \delta_3\alpha^3 + \delta_2\alpha^2 + \delta_1\alpha^1 + \alpha \tag{5}$$

The results are depicted in Fig. 3 for α varying between 0.01 and 1 and tabulated in Table 2.

Fig. 3 shows that there is a quite smooth variation of β_{opt} with excess bandwidth, α. Overall, the value β_{opt} increases with timing offset t/T for values of excess bandwidth α between 0.15 and 1.

Table 2 Polynomial Coefficients in Relation (5)

t/T	δ_3	δ_2	δ_1	δ_0
0.05	1.64103	−1.80513	1.2359	−0.0417949
0.1	4.9641	−7.59385	4.00359	−0.353846
0.2	1.7641	−3.14051	2.50359	−0.16729
0.3	0.317949	1.22103	−0.0587179	−0.410256

A different behaviour is manifested for higher values of excess bandwidth α ($\alpha > 0.55$) for t/T=0.05 and t/T=0.1.

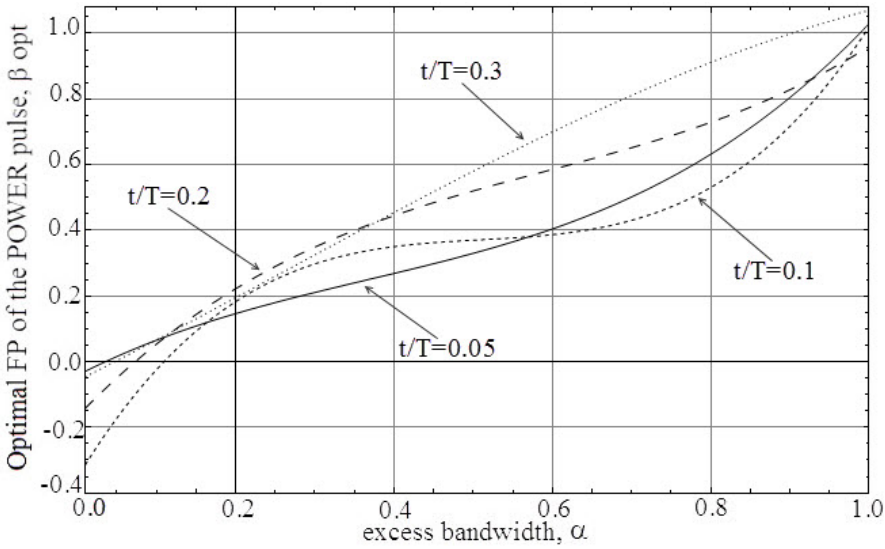


Fig. 3 Cubic relationship between the optimal FP of the POWER pulse β, and the excess bandwidth α

Also, for smaller values of excess bandwidth α ($0.12 < \alpha < 0.35$) for t/T=0.2 and t/T=0.3, the results lead us to the observation that the variation slope is higher.

Table 3 The ISI error probability, for optimal and non-optimal values of the parameter β

α	t/T=0.05		t/T=0.1		t/T=0.2	
		Pe				Pe
0.25	$\beta = 0.25$	4.576*10^-8	$\beta = 0.25$	8.243*10^-7	$\beta = 0.25$	2.048*10^-4
	β_{opt} =0.18	**4.543*10^-8**	β_{opt} =0.25	8.241*10^-7	β_{opt} **=0.29**	**2.030*10^-4**
0.35	$\beta = 0.33$	3.103*10^-8	$\beta = 0.33$	3.564*10^-7	$\beta = 0.33$	6.434*10^-5
	β_{opt} =0.24	**3.063*10^-8**	β_{opt} =0.33	3.564*10^-7	β_{opt} **=0.40**	**6.225*10^-5**
0.50	$\beta = 0.37$	1.945*10^-8	$\beta = 0.37$	1.250*10^-7	$\beta = 0.37$	1.662*10^-5
	β_{opt} =0.33	**1.938*10^-8**	β_{opt} =0.37	1.25042*10^-7	β_{opt} **=0.52**	**1.495*10^-5**
1.00	$\beta = 1.02$	1.228*10^-8	$\beta = 1.02$	2.941*10^-7	$\beta = 1.02$	1.359*10^-6
	β_{opt} =1.03	**1.227*10^-8**	β_{opt} =1.02	2.941*10^-8	β_{opt} **=0.96**	**1.345*10^-6**

In Table 3 we can now compare the performance of the POWER pulse for the optimal and non-optimal values of the parameter β.

We observe improved results for the optimal case.

4 Conclusion and Future Work

In this study we have shown the necessity of using an optimization scheme for enhanced performance in terms of bit error probability.

The bit error performance of POWER pulse was optimized for a binary symmetric channel with inter-symbol interference and additive noise using Nelder-Mead technique. Our results imply that the design parameter β, used as free parameter (FP) was found to follow a cubic relationship with the excess bandwidth factor, α.

In a future work we will intend to study the effect of pulse shaping in optical coherent systems and to establish a trade-off between the required average optical power and the bandwidth, using pulses with a minimum number of design parameters.

Acknowledgments. This work was supported by the project CNCSIS-UEFISCDI Project number RU-107/2010.

References

1. Franks, L.E.: Further results on Nyquist's problem in pulse transmission. IEEE Trans. Commun. Technol. 16, 337–340 (1968)
2. Hill Jr., F.S.: A unified approach to pulse design in data transmission. IEEE Trans. Commun. 25, 346–354 (1977)
3. Beaulieu, N.C., Tan, C.C., Damen, M.O.: A better than Nyquist pulse. IEEE Commun. Lett. 5, 367–368 (2001)
4. Beaulieu, N.C., Damen, M.O.: Parametric construction of Nyquist-I pulses. IEEE Trans. Commun. 52, 2134–2142 (2004)
5. Assalini, A., Tonello, A.M.: Improved Nyquist pulses. IEEE Commun. Lett. 8, 87–89 (2004)
6. Sandeep, P., Chandan, S., Chaturvedi, A.K.: ISI-Free pulses with reduced sensitivity to timing errors. IEEE Commun. Lett. 9, 292–294 (2005)
7. Chandan, S., Sandeep, P., Chaturvedi, A.K.: A family of ISI-free polynomial pulses. IEEE Commun. Lett. 9, 496–498 (2005)
8. Assimonis, S.D., Matthaiou, M., Karagiannidis, G.K.: Two-parameter Nyquist pulses with better performance. IEEE Commun. Lett. 12, 807–809 (2008)
9. Assimonis, S.D., Matthaiou, M., Karagiannidis, G.K., Nossek, J.A.: Parametric construction of improved Nyquist filters based on inner and outer functions. In: IEEE International Conference on Communications, Dresden, Germany, pp. 1976–1980 (June 2009)
10. Alexandru, N.D., Onofrei Balan, A.L.: Improved Nyquist filters with piece-wise parabolic frequency characteristic. IEEE Commun. Lett. 15, 473–475 (2011)
11. Assimonis, S.D., Matthaiou, M., Karagiannidis, G.K., Nossek, J.A.: Improved Parametric Families of ISI-free Nyquist Pulses Using Inner and Outer Functions. IET Signal Processing 5, 157–163 (2011)
12. Alexandru, N.D., Alexandru, M.L.: A Comparison of Double Convex and Double Concave Improved Nyquist Filters. Advances in Electrical and Computer Engineering 10, 63–66 (2010)
13. Mohri, M., Hamamura, M.: ISI-free Power Roll-Off Pulse. IEICE Trans. Fundamentals E92-A(10), 2495–2497 (2009)
14. Alexandru, N.D., Balan, A.L.: ISI-free pulses produced by improved Nyquist filter with piece-wise linear characteristic. Electronics Letters 47(4), 256–257 (2011)
15. Beaulieu, N.C.: The evaluation of error probabilities for intersymbol and cochannel interference. IEEE Trans. Commun. 31, 1740–1749 (1991)

Influence of Impulse Noise on Alamouti Code Performances

Mihaela Andrei, Lucian Trifina, and Daniela Tarniceriu

Abstract. Multiple-Input Multiple-Output systems ensure spatial diversity and fading protection for wireless communications. This paper addresses the impulse noise described by the Middleton Class-A statistical model, which can affect their performances. We have analyzed the influence of impulse noise, described by Middleton Class-A non-Gaussian model on the performances of the Alamouti code, with two transmitting antennas and two receiving antennas, on channels affected by Rayleigh fading, with Binary Phase-Shift Keying modulation. The simulations were made for different parameter values of the noise model and they showed that as the Signal-to-Noise Ratio increases, the performances decrease for the impulse noise compared with the Gaussian one. It is shown that BPSK modulation is more robust than QPSK one, used in transmission over impulse noise environment using an Alamouti code.

1 Introduction

Optimizing wireless communications entails the possibility of transmitting a large amount of information in a very short period of time and with as few errors as possible. A factor that influences significantly the communication quality is the propagation environment. This can lead to either attenuation of the received signal, known as *fading*, or introduction of certain delays, or phase shifting for one or more frequency components. All of these lead to the receiver's inability to recover the signal. The solution for this problem is to transmit more signal replicas, a technique known as *diversity* (spatial, temporal or frequency). [1]

Mihaela Andrei
Department of Electronics & Telecommunications,
"Dunarea de Jos" University of Galati, Romania
e-mail: mihaela.andrei@ugal.ro

Lucian Trifina · Daniela Tarniceriu
Department of Telecommunications,
"Gheorghe Asachi" Technical University Iasi, Romania
e-mail: {luciant,tarniced}@etti.tuiasi.ro

L. De Strycker (ed.), *ECUMICT 2014*,
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_2, © Springer International Publishing Switzerland 2014

The systems that provide spatial diversity and successfully repel the fading are the Multiple-Input Multiple-Output (MIMO) ones. They involve multiple antennas for both transmission and reception. Among the encoding techniques used for the above-mentioned channels, a special interest is given to space-time codes (block, trellis or Bell labs LAyered Space-Time - BLAST), because they improve the data transmission safety, especially at high speeds [2].

The signals are not affected only by fading, but also by noise. For the majority of the proposed codes, the channel affected by fading and Gaussian white noise (AWGN) was considered, ignoring other sources of noise, like: industrial noise, man-made activity such as automobile spark plugs [3], microwave ovens [4] and network interference [5], noises known to be non-Gaussian.

So, we need to perform an analysis of communication systems in the presence of impulse noise (non-Gaussian) and, obviously, to try to eliminate or at least diminish its disruptive effects. The Middleton Class-A model is frequently used for modeling the impulse noise.

This type of noise has been used for investigating the performances of the orthogonal space-time codes (OSTBC), for QPSK and 16QAM modulations, respectively, on MIMO channels affected by Rayleigh fading [6] by comparing the Symbol Error Rate (SER) curves with the ones obtained when only the Gaussian noise was present. By varying the parameters that describe the noise model, for low values of Signal-to-Noise Ratio (SNR), the coding gain drops with at most 6dB compared to the Gaussian model, after which, along with the SNR increases, the system's performance drops in the case of non-Gaussian noise.

Most of the systems affected by non-Gaussian noise suffer performance degradation for high SNR values [7]. For example, in [8] an increase of the Bit Error Rate (BER) is observed for IEEE 802.11a and IEEE 802.11b, under the influence of Middleton Class-A noise compared to AWGN.

This paper proposes an analysis of the non-Gaussian noise, expressed through the Middleton Class-A statistical model. It investigates the OSTBC Alamouti code performances, on channels affected by Rayleigh fading, with BPSK modulation, and maximum likelihood (ML) receiver, with non-Gaussian and AWGN, respectively. Several parameters modeling the non-Gaussian noise were considered.

The paper is structured as follows. Sect. 2 describes the Middleton Class-A impulse noise model and Sect. 3 presents the system model. The simulation results are shown in Sect. 4 and conclusions are highlighted in Sect. 5.

2 Middleton Class-A Model

The non-Gaussian noise has a Gaussian component (n_g), with variance σ_g^2, and an impulse one (n_i), with variance σ_i^2. Thereby, the model for non-Gaussian noise can be considered as Additive White Class A Noise (AWCN):

$$n = n_g + n_i, \tag{1}$$

whose probability density function follows a Middleton Class-A distribution [6]. Its expression, for complex noise, is given by:

$$p(n) = \sum_{m=0}^{\infty} \frac{A^m e^{-A}}{\pi m! \sigma_m^2} exp\left(-\frac{|n|^2}{\sigma_m^2}\right) \tag{2}$$

We can observe that this is a Poisson weighted sum of Gaussian distributions. In (2), the terms have the following meaning: m is the number of active interferences (or impulses), and A is the impulse index and indicates the average number of impulses during interference time. This parameter allows the description of noise as follows: as A gets smaller, the noise gets more impulsive; conversely, as A grows, the noise tends towards AWGN.

σ_m^2 represents the noise's total variance and it is given by:

$$\sigma_m^2 = \frac{\frac{m}{A} + T}{1 + T}, \tag{3}$$

where

$$T = \frac{\sigma_g^2}{\sigma_i^2} \tag{4}$$

is called Gaussian factor. From its expression, we can observe that, for low T values, the impulsive component dominates, and for large T values, the AWGN component is the one that prevails.

3 Mathematical Model of MIMO System

MIMO communication channels involve multiple antennas for both transmission and reception, this way achieving spatial diversity. In this paper we assume that the propagation channels are without memory and they are affected by flat fading (Rayleigh type). Let there be a system with N_T emitting and N_R receiving antennas.

During one symbol, the transmitted signals x_i, form a column array, denoted by **x**, of size [N_T, 1]:

$$\mathbf{x} = \left[x_1, x_2 \ \ldots \ x_{N_T} \right]^T , \tag{5}$$

where i represents the index of the emitting antenna.

The linear input-output relation for the MIMO channel is:

$$\mathbf{r} = \mathbf{H} \cdot \mathbf{x} + \mathbf{n} \tag{6}$$

where **r** is the array of signals received by the N_R antennas, of size [N_R, 1]:

$$\mathbf{r} = \left[r_1, r_2 \ \ldots \ r_{N_R} \right]^T \tag{7}$$

and **H**, of size [N_R x N_T], is the channel matrix, also called the transfer function, having the form at moment t:

$$\mathbf{H}_t = \begin{bmatrix} h_{1,1}^t & h_{1,2}^t & \ldots & h_{1,N_T}^t \\ h_{2,1}^t & h_{2,2}^t & \ldots & h_{2,N_T}^t \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_R,1}^t & h_{N_R,2}^t & \ldots & h_{N_R,N_T}^t \end{bmatrix} \tag{8}$$

The $h_{i,j}$ coefficients are actually the channel fading coefficients between the emitting antenna i and the receiving antenna j. These coefficients change in time and are described by various statistical models. For our study, we assume the Rayleigh model for which the fading coefficients are random complex Gaussian variables, with identical distribution with zero mean and unit variance.

n from (6) is the column array of noise (Gaussian or impulsive):

$$\mathbf{n} = \left[\eta_1, \eta_2 \ \ldots \ \eta_{N_R} \right]^T \tag{9}$$

At moment t, the signal received by antenna j will be given by:

$$r_j^t = \sum_{i=1}^{N_T} h_{ji}^t \cdot x_i^t + \eta_j^t , \tag{10}$$

For the specific scheme proposed by Alamouti, there are two emitting antennas and N_R receiving ones. In this paper, we will consider two emitting and two receiving antennas, with BPSK modulation. The advantage of the Alamouti space-time codes consists in developing space-time diversity and decoding [9].

The encoder structure proposed by Alamouti is shown in fig. 1:

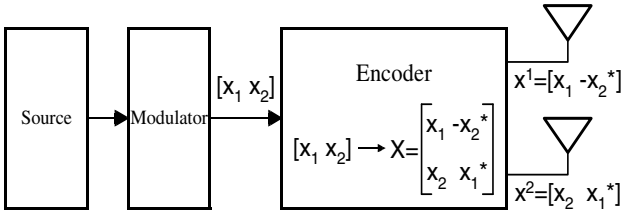


Fig. 1 Alamouti encoder structure

At each coding operation, the group of the two modulated symbols is transmitted according to the following scheme:

$$\begin{array}{c|cc} & t & t+\tau \\ \hline T_{x1} & x_1 & -x_2^* \\ T_{x2} & x_2 & x_1^* \end{array} \tag{11}$$

At moment t, the first antenna (denoted by T_{x1}) transmits the signal x_1, and the T_{x2} antenna, the signal x_2. At the following moment, $t+\tau$, the signals emitted by the two antennas are $-x_2^*$ and x_1^*, respectively.

At the receiving end, the signals are given by:

$$\begin{cases} r_{j,1} = h_{j,1} \cdot x_1 + h_{j,2} \cdot x_2 + \eta_{j,1} \\ r_{j,2} = -h_{j,1} \cdot x_2^* + h_{j,2} \cdot x_1^* + \eta_{j,2} \end{cases} \tag{12}$$

The matrix form is:

$$\mathbf{r}_j = \begin{bmatrix} r_{j,1} & r_{j,2} \end{bmatrix} = \begin{bmatrix} h_{j,1} & h_{j,2} \end{bmatrix} \begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix} + \begin{bmatrix} \eta_{j,1} & \eta_{j,2} \end{bmatrix} \tag{13}$$

The decoding is based on the maximum likelihood algorithm, which selects the most probable symbols $\hat{x}_1$ and $\hat{x}_2$. Considering that the information source is without memory, the modulated symbols x_2 and x_1 are independent from each other. Hence, separate decoding of the two symbols is possible:

$$\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} \underset{\hat{x}_1 \in S}{arg\,min} \left(\sum_{j=1}^{N_R} \left| \tilde{r}_{j,1} - \left(\left| h_{j,1} \right|^2 + \left| h_{j,2} \right|^2 \cdot \hat{x}_1 \right) \right|^2 \right) \\ \underset{\hat{x}_2 \in S}{arg\,min} \left(\sum_{j=1}^{N_R} \left| \tilde{r}_{j,2} - \left(\left| h_{j,1} \right|^2 + \left| h_{j,2} \right|^2 \cdot \hat{x}_2 \right) \right|^2 \right) \end{bmatrix}, \tag{14}$$

where

$$\tilde{\mathbf{r}}_j = \begin{bmatrix} \tilde{r}_{j,1} & \tilde{r}_{j,2} \end{bmatrix}^T = \mathbf{H}^H \cdot \mathbf{r}_j^T \tag{15}$$

and $(.)^H$ stands for conjugate transpose of the matrix.

4 Simulation Results

In this section, firstly we analyze the Middleton Class-A impulse noise, for various parameters that describe its statistical model and then we investigate its influence on the Alamouti code performances for two emitting and two receiving antennas.

4.1 Pdf Analysis

For the impulse noise analysis, we considered the model described in [6]. The simulations were done in Matlab for a number of 10^4 samples, by varying the model's A and T parameters. The Middleton Class-A noise was generated by the InterferenceModeling and Mitigation Toolbox [10]. The values for A were considered in the range $[10^{-4}, 1]$, and for T in the range $[10^{-2}, 1]$.

Figure 2 shows a comparison between the normal distribution with zero mean and unit variance and the Middleton Class-A model, for different combinations of parameters A=1, 0.1 and T=1, 0.1. We can observe that along with decreasing the impulse index A, the AWCN noise's distribution is "narrower" and taller than the Gaussian one and, when T grows, the AWCN distribution approaches the normal one.

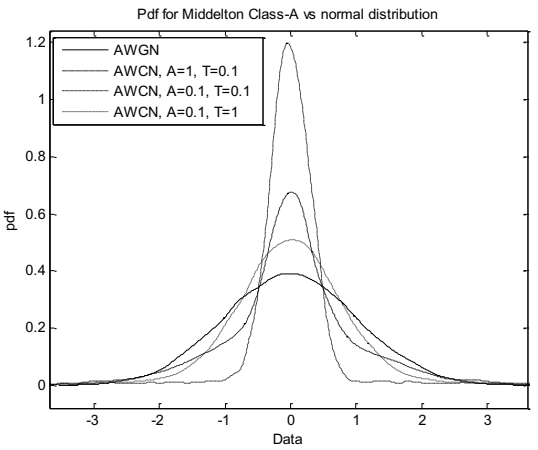


Fig. 2 Middelton Class-A distribution vs normal distribution

For A=1 and T=1, the two distributions are identical. To emphasize the impulse noise deviation from the normal distribution, we chose the parameters A=0.1 and

T=0.1. In Figure 3, this deviation is exemplified with the help of the cumulative density function.

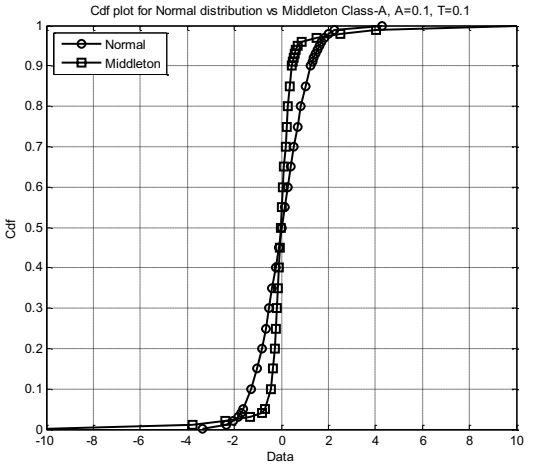


Fig. 3 Middleton Class-A samples distribution vs normal distribution

Figure 4 shows two probability density functions: the real one, obtained by means of generated noise samples and the estimated one, obtained on the base of a normal kernel function, using a window parameter that is a function of the number of points in data samples. The distribution parameters are A=0.1 and T=0.1.

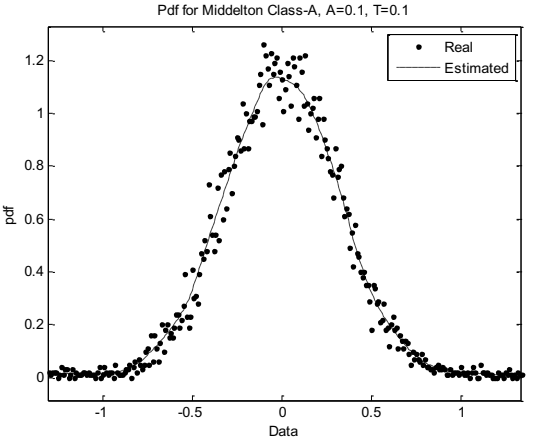


Fig. 4 Pdf for Middelton Class-A, A=0.1, T=0.1

Figure 5 shows the estimated pdfs for the Middleton Class-A noise, for a fixed value T=0.1 of the Gaussian factor and different impulse index values $A \in [10^{-4}, 1]$.

An interesting aspect that can be observed is that for a very small value of A, A=0.0001, the distribution approaches the normal one. This happens because the impulses are very rare or even singular, but with high amplitude, in order to have the same power. Therefore the distribution is practically close to a normal one, the noise being in this case predominantly Gaussian.

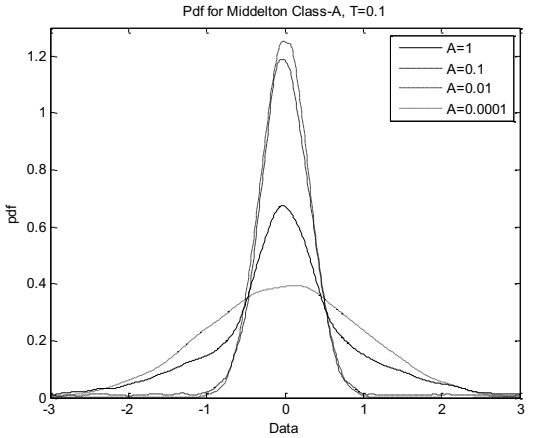


Fig. 5 Estimated pdfs for Middleton Class-A, T=0.1

In Figure 6, the parameter A was fixed at 0.01 and the Gaussian factor T was the one being varied. This influenced the Middleton Class-A distribution by considerably increasing the peak value of the pdf, compared to the normal one.

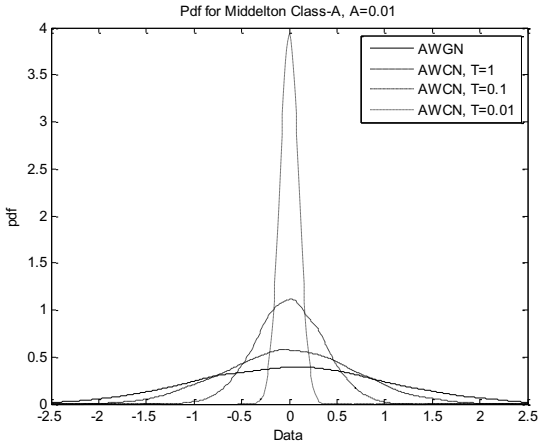


Fig. 6 Estimated pdf for Middleton Class-A, A=0.01

4.2 Bit Error Rate Analysis

Figures 7 and 8 show the Alamouti code performances using the Middleton Class-A type of impulsive noise, with two emitting and two receiving antennas, for a channel affected by Rayleigh fading and BPSK modulation. The code performance is evaluated through the Bit Error Rate (BER) curves, for various values of the parameters A and T, respectively. The simulations were done for A=0.01, 0.1, 1 and T=0.01 fixed; and for T=0.01, 0.1, 1 and A=0.01 fixed, respectively. The analysis is performed by reference to the results in [6] We mention that the performances from [6] of the same Alamouti code, evaluated through the Symbol Error Rate (SER), are given in figure 2, for the QPSK modulations, 16-QAM and for A=0.01, 0.1, 1 and T=0.01 fixed. Because for a BPSK modulation there is a single bit per symbol, this means that BER and SER are identical. Therefore, we can compare the results we obtained with those from [6] in terms of SER.

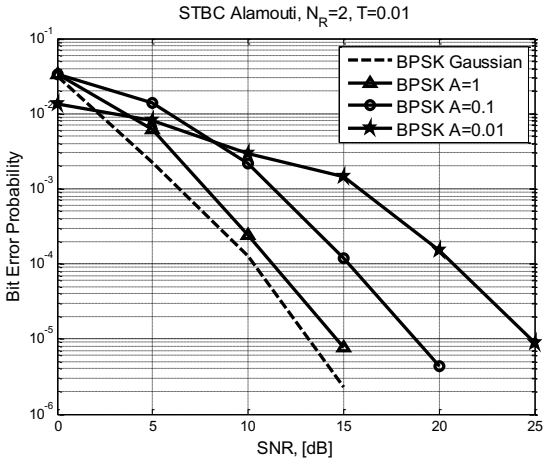


Fig. 7 Performance of Alamouti code, for different impulse index A and fixed T

For AWGN noise, the Alamouti code ensures an encoding gain of approximately 11dB [11]. Figure 7 shows the fact that, in the presence of impulse noise, the system has weaker performances than in the case of Gaussian noise. Starting from SNR=8dB, once the A parameter is lowered, the BER increases, the poorest results being for A=0.01. For low SNR values, the system behaves better in the presence of impulsive noise, for A=0.01. As the SNR increases, the performances drop significantly for the Middleton Class-A model, with A=0.01, compared to AWGN. This happens because, at high SNR values, the impulsive component has a significant influence. Comparing the code performances for different values of A, we observe that for high values of A, the BER is lower, and for A=1, it approaches AWGN.

Unlike the results in [6] for QPSK modulation, we can observe that for SER=10^{-4} and A=0.01, in the case of BPSK modulation, an approximately 4dB lower SNR is needed; when A=0.1, SNR is smaller with approximately 3 dB, and

when A=1, SNR is smaller with approximately 3.5 dB. For AWGN only, the difference between SNRs for QPSK and BPSK is approximately 5 dB. For low SNR values, when A=0.01, the BPSK modulation leads to better performances than QPSK. For example, at SNR=0 dB, in the case of QPSK modulation, SER=$3\text{x}10^{-2}$, compared to the BPSK modulation, when SER=10^{-2}. These performance differences were expected, because the QPSK modulation is more sensitive to noise than the BPSK modulation, the minimum distance between the QPSK constellation points being $\sqrt{2}$, compared to 2 for BPSK.

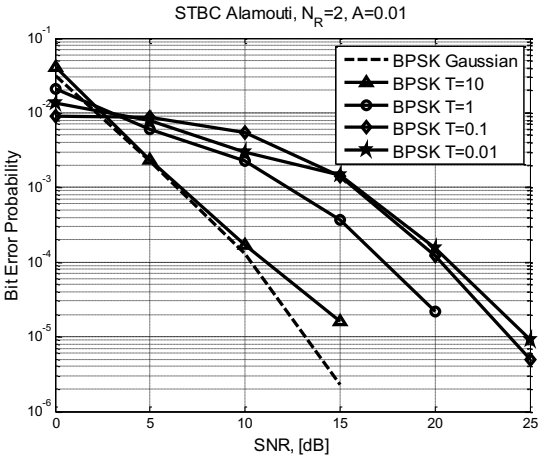


Fig. 8 Performance of Alamouti code, for different T and fixed A

In Figure 8, the parameter A was considered to be constant, at 0.01, and the parameter T was varied, in order to observe its influence on the code's performances. The same conclusion can also be drawn in this case: for SNR values higher than 4dB, when we decrease the Gaussian factor T, the performances drop compared to the AWGN case. For T=1, the BER values are very close to AWGN, and for T=0.1, respectively 0.01, they are different from AWGN, but almost identical to each other. For low SNR values, below 4dB, the opposite occurs: performances slightly increase as T decreases. We cannot anymore compare the obtained results with those in [6], because we have used two receiving antennas, unlike the case addressed in [6], with one receiving antenna.

5 Conclusions

To model the impulsive noise, a Middleton Class-A model was used. A comparative analysis between this type of noise and AWGN was conducted using the probability distributions. The conclusion drawn was that by varying the parameters that describe the model and for lower values of the impulse index and Gaussian factor, the distribution significantly differs from the normal one, except

for A=0.0001. The influence of the Middleton Class-A noise on the Alamouti code performances was investigated, using two emitting and two receiving antennas, for a channel affected by Rayleigh fading and BPSK modulation. Simulations shown that, for SNR values above 8dB, the performances drop considerably, compared to the Gaussian noise, as the impulse model's parameters get lower. For A=1 or above and for T>1, the BER values almost reach the ones obtained for the AWGN case. For smaller SNR values, the performances improve in the case of impulsive noise for parameter values as low as possible. The BPSK modulation is more robust than QPSK one, used in transmission over impulse noise environment using an Alamouti code.

References

1. Guey, J.-C., Fitz, M.P., Bell, M.R., Kuo, W.-Y.: Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels. In: Proc. IEEE VTC 1996, pp. 136–140 (1996)
2. Tarokh, V., Seshadri, N., Calderbank, A.R.: Space–time codes for high data rate wireless communication: Performance analysis and code construction. IEEE Transactions on Information Theory 44(2), 744–765 (1998)
3. Middleton, D.: Statistical-physical models of electromagnetic interference. IEEE Trans. Electromagn. Compat. EMC-19(3), 106–127 (1977)
4. Kanemoto, H., Miyamoto, S., Morinaga, N.: Statistical model of microwave oven interference and optimum reception. In: Proc. IEEE ICC 1998, pp. 1660–1664 (October 1998)
5. Win, M., Pinto, P., Shepp, L.: A mathematical theory of network interference and its applications. Proc. IEEE 97(2), 205–230 (2009)
6. Gong, Y., Wang, X., He, R., Pang, F.: Performance of Space-Time Block Coding under Impulsive noise Enviroment. In: Proc. IEEE of 2nd International Conference on Advanced Computer Control, vol. 4, pp. 445–448 (March 2010)
7. Madi, G., Sacuto, F., Vrigenau, B., Agba, B.L., Pousser, Y., Vauzelle, R., Gagnon, F.: Impacts of impulsive noise from partial discharges on wireless systems performance: application to MIMO precoders. EURASIP Journal on Wireless Communications and Networking (2011)
8. Bhatti, S.A., Shan, Q., Glover, I.A., Atkinson, R., Portugues, I.E., Moore, P.J., Rutherford, R.: Impulsive noise modeling and prediction of its impact on the performance of WLAN receiver. In: 17th European Signal Processing Conference (EUSIPCO 2009), pp. 1680–1684 (August 2009)
9. Alamouti, S.M.: A simple transmit diversity technique for wireless communications. IEEE Journal on Selected Areas in Comm 16(8), 1451–1458 (1998)
10. Gulati, K., Nassar, M., Chopra, A., Ben Okafor, N., DeYoung, M., Aghasadeghi, N., Sujeeth, A., Evans, B.L.: InterferenceModeling and Mitigation Toolbox 1.6, for Matlab. ESP Laboratory, ECE Dept., Univ. of Texas at Austin (October 2001)
11. Jafarkani, H.: Space-Time Coding Theory and Practice, p. 58. Cambridge University Press (2005)

Making OpenID Mobile and Privacy-Friendly

Faysal Boukayoua, Karel Dewitte, and Vincent Naessens

Abstract. OpenID is a widely used single sign-on standard that allows users to access different services using the same authentication. However, its usage poses a number of issues regarding privacy and security. This paper evaluates the OpenID standard and introduces three mobile strategies, two of which are validated using a prototype implementation. Significant privacy and trust improvements are attained through the use of an identity management architecture that leverages the properties of a tamperproof module. Furthermore, our approach makes OpenID more suitable for omnipresent mobile use. We remain interoperable with the OpenID standard and no modifications to the mobile platform are required.

1 Introduction

It goes without saying that Web services have become commonplaces of everyday life. Their mobile use has soared during the past years. Authentication remains an important concern throughout this evolution. A well-established and popular standard is OpenID. Many online parties provide their users with proprietary

Faysal Boukayoua · Karel Dewitte · Vincent Naessens
KU Leuven, Department of Computer Science, Technology Campus Ghent,
Gebroeders De Smetstraat 1, 9000 Ghent, Belgium
e-mail: {faysal.boukayoua,vincent.naessens}@cs.kuleuven.be,
karel.dewitte@hotmail.com

L. De Strycker (ed.), *ECUMICT 2014,*
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_3, © Springer International Publishing Switzerland 2014

authentication accounts that are OpenID-enabled. This frees users from having to maintain different accounts and they only have to sign in once to access services that support OpenID.

Boukayoua et al [2] outline a number of complementary weaknesses and strengths in network-based and claim-based identity management systems. In a network-based system, storage of user attributes is centralised at the identity provider, making it a high-value attack target. In many such systems, this actor is also directly involved in the authentication, increasing its capabilities to monitor, link and profile user transactions. On the other hand, network-based systems are typically more standardised than claim-based ones and they also require less modification to workstations. The authors present a hybrid architecture to leverage the strengths of both systems and to mitigate their weaknesses. It is validated in this work with OpenID as the network-based part and the architecture by Vossaert et al [16] as the claim-based part.

Contribution: This paper addresses multiple privacy and security issues in OpenID. We present three strategies to make it more secure and privacy-friendly. The three are compared in terms of privacy, trust, security and interoperability properties. Two of the three strategies are validated with a prototype implementation. Additionally, OpenID is made more suitable for today's authentication needs, as all three approaches incorporate commodity mobile devices.

The remainder of this paper is structured as follows. Section 2 provides further background and gives an overview of related work. The integration strategies are put forward in section 3. The prototype implementations are described in section 4, while section 5 presents an evaluation and comparison of the three approaches. Finally, conclusions are drawn in section 6.

2 Background and Related Work

OpenID is a network-based single sign-on protocol. It allows the user to prove ownership of an identifier, typically a URL. Version 2 of the standard was made public in 2007 [9].

To start authenticating, the user provides his identifier to the relying party (1). Using this identifier, the relying party carries out a discovery, the result of which is the identity provider's URL (2). Next, a session is established between the relying party and the identity provider (3). The *association* that OpenID provides here, is insufficient. As the standard does not mandate the use of TLS[1], it is left vulnerable to man-in-the-middle attacks [7]. In a next step, the user is redirected to the identity provider (4), where he is authenticated (5). In case of success, he is supplied with an assertion from the identity provider and redirected back to the relying party (6). In the assertion, the identity provider provisions any required attributes and endorses

[1] Transport Layer Security: the successor of SSL.

the user's ownership of the given identifier. If it is authorised by the relying party (7), the user can now access the protected resource (8).

Being a standard, OpenID is well-established and widely deployed. However, its use incurs several security issues. Many OpenID providers resort to passwords, which are often reused and chosen weak [17, 15]. An exception is myOpenID[2], which offers certificate-based authentication. Furthermore, OpenID is very prone to phishing. Instead of redirecting the user to his identity provider, a malicious relying party can easily present a self-constructed page, where login info is captured [4, 7]. In addition, since an OpenID account is used for single sign-on, a permanently authenticated session with the identity provider is not uncommon. If the latter does not implement proper user authorisation, cross-site request forgery (CSRF) allows an attacker[3] to authenticate to other relying parties on behalf of the user [4, 1]. Once logged in, further exploitation could allow the adversary to perform privileged actions. Additionally, malware on the workstation can intercept the user name and password. What's more, multiple security-critical responsibilities are fulfilled by the identity provider: authentication, attribute storage, as well as provisioning. This makes it a high-value attack target [2].

OpenID performs poorly in terms of trust properties [7]. The identity provider is not required to be a trusted party, as opposed to, f.i., Shibboleth [14]. It could as well be under the user's control. Proving ownership of an identifier is by no means equivalent to making assertions about a user's identity! In the first years of OpenID, the user could provide any identifier, a practice still in use. Later, many relying parties restricted this to a handful of trusted identity providers (Google, Facebook,...).

OpenID exhibits a number of privacy issues. Typically, a globally unique, DNS-based identifier is used, making user transactions linkable by relying parties. In addition, the identity provider's intervention is required for each authentication, giving it pervasive capabilities for monitoring, linking and profiling [5, 7]. Since the identity provider and the relying party are in direct contact, this allows them to collude against the user. Furthermore, many implementations are not equipped to show user feedback and to ask for consent. Others, like Google, display a feedback and consent prompt only upon first sign-on to a relying party. To the best of our knowledge, there is not yet an OpenID version that offers selective attribute disclosure, similar to what Shibboleth's uApprove [10] does.

This work considers OpenID in combination with the identity management architecture (onward referred to as the *IdM architecture*) by Vossaert et al [16]. It exploits the properties of a tamperproof module to enforce trust. A strict separation between claim providers[4] and relying parties is maintained, as their communication is mediated through the tamperproof module. The middleware displays feedback about the ongoing transaction and allows the user to consent, which includes entering his PIN. Identifiable, pseudonymous as well as (accountable) anonymous transactions are supported. Standalone mobile use of this architecture was previously validated

[2] `https://www.myopenid.com`

[3] This can be a malicious website or an attacker infecting a trusted website.

[4] To avoid confusion, claim-based identity providers are referred to as claim providers.

by Boukayoua et al [3]. However, real-life use requires an approach that is more interoperable with existing infrastructure.

Related research is presented by Dodson et al. [8], proposing a phone-based, asymmetric challenge-response protocol for user authentication to the OpenID provider. The approach by Feld et al [11] is similar, which combines the German eID with OpenID. Both approaches offer protection against phishing and replace user names and passwords by strong authentication. However, they do not propose a trust infrastructure to endorse users' identity information. In addition, the OpenID provider still stores user information, thus remaining a high-value attack target. Leicher et al. [13] integrate the infrastructure of a mobile phone operator into OpenID. Added to the improvements of the two previous works, this approach introduces trust in OpenID by relying on the operator's infrastructure and on the tamperproofness of a SIM card. They also extend OpenID with support for pseudonymity and anonymity. However, the focus mainly lies on how to combine OpenID with the infrastructure of a mobile network operator.

3 Approach

To resolve the OpenID issues that are listed in section 2, three different approaches are introduced. First, we list the requirements we aim to satisfy.

3.1 Requirements

The security requirements are as follows:

- S_1 Passwords are substituted for stronger credentials.
- S_2 Protection against phishing is provided.
- S_3 Authentication-related CSRF attacks are prevented.
- S_4 The attack value and exposure of the identity provider is reduced.
- S_5 Data authentication and confidentiality is ensured between the communicating parties.

To address the privacy issues, we require the following:

- P_1 The identity provider's capabilities for transaction monitoring, linking and profiling, are reduced.
- P_2 Collusion between identity providers and relying parties is obstructed.
- P_3 The identity provider cannot impersonate the user.
- P_4 The user can selectively disclose or withhold his attributes.

As for trust, we impose that:

- T_1 relying parties can corroborate provisioned user attributes.

T_2 the required trust in the workstation is reduced, as this can be a publicly used computer.

The remaining requirements pertain to interoperability and usability:

R_1 We remain compliant with the OpenID standard.
R_2 No modifications are applied to the relying party's infrastructure.
R_3 The relying party is agnostic to the claim-based credential technology in use.
R_4 The solution is portable across workstations.
R_5 Feedback about the ongoing transaction is presented to the user and his consent is enforced.

3.2 Improvement Strategies

Each approach shifts a gradually increasing part of the identity provider's responsibilities towards the user's smartphone and/or to the tamperproof module.

3.2.1 Strategy 1: Authentication via the Mobile Device, Attribute Storage and Provisioning by the OpenID Identity Provider

In this setup, two major changes are applied to a password-based OpenID implementation. First, a mobile device with a tamperproof module is introduced. The latter contains the mobile part of the IdM architecture. It is addressed by the phone through a middleware layer. Second, password-based authentication is substituted for the one used in the IdM architecture, Elliptic Curve Diffie-Hellman. The user is authenticated out of band and his consent is first needed to proceed. Note that the identity provider still stores and provisions user attributes.

As a result, the - typically password-based - authentication in the standard protocol, takes place as follows. Using a short-range communication protocol, a session challenge is advertised to the authentication app on the user's mobile device. The app displays information about the identity provider[5] and asks for the user's permission to proceed. Upon confirmation, the challenge is sent to the tamperproof module. Prior to setting up a secure channel with the identity provider, it prompts the user to enter his PIN. The authentication response is sent over this channel. After this authentication response is verified, the standard OpenID protocol flow is resumed.

As mentioned in section 2, OpenID is vulnerable to phishing attacks by the relying party. To mitigate this, our authentication app expects the user to preconfigure a list of trusted identity providers.

[5] From the viewpoint of the IdM architecture, the OpenID identity provider is essentially a relying party, in this case requiring authentication without attribute provisioning.

3.2.2 Strategy 2: Authentication and Attribute Provisioning by a Tamperproof Module on the Device

The concerns and steps of this approach are largely similar to 3.2.1. Here however, attribute storage and provisioning are also moved to the IdM architecture. Therefore, the remaining responsibility of the identity provider is to convert claims to OpenID assertions that are understood by the relying party. The user still authenticates to the identity provider. However, he must now also agree to disclose attributes to the relying party and, if yes, indicate which ones to release. These are transferred over the secure channel that is set up upon authentication [16]. Consequently, the attributes the identity provider requires, are part of the challenge that is advertised to the mobile device. They are displayed as feedback to the user, alongside the relying party to which info is about to be disclosed. Note that the identity provider does not create the OpenID assertion using information from its own data storage. Instead, it uses the attributes that are provisioned by the tamperproof module.

3.2.3 Strategy 3: The Mobile Device Running an Entire OpenID Identity Provider

In this setup, the identity provider's functionality is entirely moved to the user's mobile and tamperproof module (mIdP). It consists of two components: a lightweight Web server and an authentication app. This app addresses the tamperproof module for security- and trust-sensitive operations. The resulting protocol is as follows. The user requests a protected resource, upon which the relying party asks for and receives his OpenID identifier. The relying party then obtains the mIdP's address through the discovery step. Subsequently, the mIdP and the relying party set up a shared session over a secure channel. The browser is then redirected to the mIdP and the user is prompted on the mobile device to review and agree to the attributes to provision and the relying party to disclose them to. Following his user's consent, the challenge and the list of required attributes are sent to the tamperproof module. The latter prompts the user to enter his PIN. If successful, an authentication and assertion response is generated. The mIdP then redirects the user's browser to the relying party and sends along the generated response. If it is correctly verified, the protected resource is made available to the user.

Note that this approach authenticates the tamperproof module in step 5: the IdM architecture's common key [16] is used to establish a TLS session between the relying party and the mIdP. Also note that no short-range communication is required to transfer the authentication challenge from the browser to the mobile device.

3.3 Trust Relations and Enforcement

As section 2 points out, OpenID falls short regarding trust establishment. On the other hand, many corporations, governments and telcos already have trust infrastructures in place. Reusing them is both cost-effective and an added value to users and relying parties.

In strategy 1, the identity provider needs to establish trust in the claim providers and the issuer of the tamperproof module. If all three fall under the same organisation, they are likely part of the same PKI[6]. If not, the identity provider must implement a black- or whitelisting mechanism. Trust is enforced by the tamperproof module and the secure channel that it maintains towards the identity provider [16]. The relying party trusts the user authentication by the identity provider. This trust is enforced by (1) TLS for direct communication between the identity provider and the relying party (2) MAC[7] authentication for communication that is mediated by the browser. The MAC key is part of the pre-established session between the relying party and the identity provider.

Similar considerations apply to the second strategy. In addition, since the identity provider now acts as an attribute mediator between the tamperproof module and the relying party, the latter needs to express which claim providers it trusts. This trust preference can be indicated in the OpenID authentication request. Alternatively the relying party can accept all claim providers that are also trusted by the identity provider. The first approach allows fine-grained control, while the second is realisable without cooperation from the identity provider.

In the third strategy, the relying party needs to establish trust in each mobile identity provider, i.e. in the issuers of its tamperproof module and in the credentials that are stored on it. Additionally, the IdM architecture can also restrict the disclosed attributes per relying party [16].

4 Proof of Concept

This section describes the proof of concept, in which strategies 1 and 3 are validated. Only the infrastructure of the user and the identity provider are addressed, as the relying party is not modified (see section 3.1).

4.1 Strategy 1

The identity provider runs in a lightweight Jetty Servlet container, both suitable for embedded devices and large-scale computing applications. The OpenID

[6] Public Key Infrastructure.
[7] Message Authentication Code.

implementation is OpenID4Java. Memcached, an in-memory key-value store, maintains session statuses between the browser and the identity provider. A session is marked as *authenticated* upon successful sign-on.

QR codes are used for short-range communication between the browser and the mobile device. These are two-dimensional barcodes that can contain up to 3kB of data. They offer multiple advantages. No extra hardware or configuration is required, as opposed to Bluetooth. The requirement of a computer screen and a smartphone camera are essentially no limitation. As authentication takes place out of band, a mechanism is required to resume the browser's workflow. A status checking page is created for this purpose. The QR authentication page contacts it, using an asynchronous long-polling request[8]: the artificially long timeout allows the smartphone authentication to complete before returning an HTTP status - 200 in case of success. To make the session known to the phone, the QR code contains its identifier. This is, however, an alternate id, to obstruct shoulder surfing. The mapping between both ids is only known to the identity provider.

The mobile device is an Acer Liquid Glow with Android 4. Smartphones are ideal for authentication, considering their widespread presence and personal nature. As a tamperproof module, we use a Giesecke & Devrient Mobile Security Card SE 1.0, running JavaCard 2.2.2. An authentication app reads the QR code using the smartphone camera, extracts the identity provider's information and displays it to the user for reviewing. The tamperproof module generates a response and sends it over a secure channel to the identity provider.

4.2 Strategy 3

In this strategy, the identity provider entirely resides on the tamperproof module and on the mobile device. Feedback and consent, communication with the tamperproof module and the authentication and attribute provisioning servlets are all confined to one Android application. This is preferred over a multi-app approach, as it poses less performance overhead. More importantly, it avoids inter-process communication and its related security concerns [6]. Additionally, it is also more portable to platforms with limited capabilities for inter-process communication. The same phone and tamperproof module as in strategy 1 are used. Note that no short-range communication is required.

TLS sessions use the card-specific common key [16] to authenticate the mobile identity provider to the relying party. The Mobile Security Card provides private key operations through a PKCS#11 interface [12]. Unfortunately, the card's processing power is not yet sufficient to handle the exchange of all OpenID messages with the relying party over an end-to-end channel. Therefore, message generation and parsing takes place on the phone rather than on the tamperproof module. We expect

[8] Repeated polling incurs greater networking overhead and HTTP WebSockets are not supported by older browsers.

this hurdle to be resolved soon, as new and more powerful tamperproof modules find their way to the market.

5 Evaluation

5.1 Requirements Evaluation

This section evaluates how well the requirements from section 3.1 are satisfied. An overview is provided in table 1. All three strategies substitute passwords for strong authentication (S1). Furthermore, as strategy 1 and 2 expect the user to preconfigure a trusted list of identity providers, phishing on the workstation is obstructed (S2). As for CSRF attacks, strategy 3 entails no session between the browser and the mobile identity provider. The user's authorisation is requested for every transaction, thus preventing authentication-related CSRF attacks (S3). While attributes are still stored by the identity provider in strategy 1, they are decentralised in strategies 2 and 3. This reduces the exposure and attack value of the identity provider (S4). As for data authentication and confidentiality, either TLS or the IdM architecture's end-to-end secure channel [16] is used between every two communicating parties (S5). Strategies 1 and 2 do not protect against monitoring, linking and profiling by the identity provider, which is directly involved in every transaction. Strategy 3 resolves this by placing the entire identity provider on the user's mobile device and tamperproof module (P1-P3). Furthermore, the user is provided with the means to review and consent to each transaction (P4). Section 3.3 explains how to reuse existing infrastructures to make OpenID trust-enabled (T1). Additionally, none of the three approaches involves entering credentials on the workstation (T2). This is attained at the expense of requiring more trust in the mobile device. This is, however, a reasonable choice, considering the more elaborate security model of today's mobile platforms and the use of a tamperproof module. Also, the user is likely to trust his mobile device more than a public computer for authentication.

As for the other requirements, none of the approaches modify the OpenID standard (R1). Furthermore, the relying party's infrastructure does not need to change (R2), with the exception of optionally maintaining black- or whitelists of claim providers and/or tamperproof module issuers (see section 3.3). Furthermore, the relying party is not exposed to details of any claim-based credential technology (R3). The shift towards mobile devices, makes the three strategies portable across workstations (R4). Lastly, all three also use the mobile device to display feedback about the ongoing authentication (R5).

Table 1 Comparison of how well each strategy satisfies the privacy, security and trust requirements (RP = relying party, IdP = identity provider)

	Password-based OpenID	**Strategy 1**	**Strategy 2**	**Strategy 3**
Strong authentication instead of passwords	IdP-dependent	Yes	Yes	Yes
Phishing protection	No	Yes	Yes	Yes
Protection against authN-based CSRF attacks	IdP-dependent	IdP-dependent	IdP-dependent	Yes
Reduced attack value of IdP	No	No	Yes	Yes
Data authN and confidentiality between parties	IdP- and RP-dependent	RP-dependent	RP-dependent	RP-dependent
Reduced linking monitoring, profiling by IdP	No	No	No	Yes
Prevent collusion between IdP and RP	No (IdP and RP communicate directly)			Yes
User impersonation by IdP prevented	No	No	No	Yes
Selective attribute disclosure	IdP-dependent	IdP-dependent	Yes	Yes
Reduced trust in workstation	No	Yes	Yes	Yes

5.2 Network Anonymity

Mobile devices are usually tied to an individual user. Therefore, strategy 3 may lead to increased network anonymity concerns: a mobile identity provider with a publicly unique IP address, allows the relying party to link transactions. Approaches to mitigate this, must ensure that the mobile identity provider remains reachable as a Web service. The simplest approach is for the user to rely on a frequently changing IP address. However, this is rarely under his control. Alternatively, a reverse proxy service can (1) be shared by multiple devices (2) generate random, changing subdomains or URL segments that correspond to each mobile identity provider. However, this requires trust in the party that operates this proxy. Another possibility is for the mobile identity provider to register as a hidden service on anonymity networks like Tor or I^2P. Typical use, however, involves both the relying party and the mobile identity provider being connected to such a network. In the case of Tor, this can be mitigated by using Tor2web. This is a widely hosted WWW-to-Tor gateway software, allowing the exposure of anonymised websites to clients that are not running Tor.

5.3 Discussion

The interoperability and usability requirements are fulfilled by all three strategies. Furthermore, a tradeoff can be observed in terms of infrastructural prerequisites and

privacy and security gains. The merits of strategy 1 are secure authentication and decreased trust in the workstation. In return, a mobile device with the IdM architecture is needed. The privacy benefits are limited. Strategy 2 moves attribute provisioning to the mobile device. This reduces the attack value of the OpenID provider. It also also allows the user to selectively disclose his attributes. However, this strategy entails far-reaching modifications to the identity provider, in which significant trust is still required. Strategy 3 offers the best privacy properties. It does away completely with the OpenID provider as a third party. As a result, the user has full control over his transactions and over the attributes he discloses. Here however, the mobile device must be reachable as an online server. Additionally, network anonymity becomes an important concern, as mobile devices are typically personal.

6 Conclusion

This paper has presented three hybrid strategies to address multiple security and privacy issues in OpenID. Properties like trust and interoperability are also taken into account. Through a comparative evaluation, we have shown significantly better performance -each strategy to its own extent- than existing OpenID implementations. The use of commodity mobile devices makes OpenID more suitable for today's authentication needs. No modifications to OpenID nor the mobile platform are needed. Two of the three strategies are validated by a prototype.

Acknowledgements. This work was made possible through funding from the *MobCom* and *SecureApps* projects, granted by the Flemish *agency for Innovation by Science and Technology (IWT)*.

References

1. Barth, A., Jackson, C., Mitchell, J.C.: Robust defenses for cross-site request forgery. In: Proceedings of the 15th ACM Conference on Computer and Communications Security, CCS 2008, pp. 75–88. ACM, New York (2008)
2. Boukayoua, F., Vossaert, J., De Decker, B., Naessens, V.: Claim-based versus network-based identity management: A hybrid approach. In: Schmidt, A.U., Russello, G., Krontiris, I., Lian, S. (eds.) MobiSec 2012. LNICST, vol. 107, pp. 38–50. Springer, Heidelberg (2012)
3. Boukayoua, F., Vossaert, J., De Decker, B., Naessens, V.: Using a smartphone to access personalized web services on a workstation. In: Camenisch, J., Crispo, B., Fischer-Hübner, S., Leenes, R., Russello, G. (eds.) Privacy and Identity 2011. IFIP AICT, vol. 375, pp. 144–156. Springer, Heidelberg (2012)
4. Brands, S.A.: The problem(s) with openid (August 2007), `http://untrusted.ca/cache/openid.html`
5. Chadwick, D.W.: Federated identity management. In: Aldini, A., Barthe, G., Gorrieri, R. (eds.) FOSAD 2007/2008/2009. LNCS, vol. 5705, pp. 96–120. Springer, Heidelberg (2009)

6. Chin, E., Felt, A.P., Greenwood, K., Wagner, D.: Analyzing inter-application communication in android. In: Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services, MobiSys 2011, pp. 239–252. ACM, New York (2011)
7. van Delft, B., Oostdijk, M.: A security analysis of openid. In: de Leeuw, E., Fischer-Hübner, S., Fritsch, L. (eds.) IDMAN 2010. IFIP AICT, vol. 343, pp. 73–84. Springer, Heidelberg (2010)
8. Dodson, B., Sengupta, D., Boneh, D., Lam, M.S.: Secure, consumer-friendly web authentication and payments with a phone. In: Conference on Mobile Computing, Applications, and Services, MobiCASE 2010, Santa Clara, CA, USA (2010)
9. Fitzpatrick, et al.: OpenID authentication 2.0 - final (December 2007)
10. Orawiwattanakul, et al.: User-controlled privacy protection with attribute-filter mechanism for a federated sso environment using shibboleth. In: Proceedings of 3PGCIC 2010, pp. 243–249. IEEE Computer Society, Washington, DC (2010)
11. Feld, S., Pohlmann, N.: Security analysis of openid, followed by a reference implementation of an npa-based openid provider. In: ISSE 2010 Securing Electronic Business Processes, pp. 13–25. Vieweg+Teubner (2011)
12. RSA Laboratories. Pkcs #11 v2.30: Cryptographic token interface standard (April 2009), http://www.emc.com/emc-plus/rsa-labs/standards-initiatives/pkcs-11-cryptographic-token-interface-standard.htm
13. Leicher, A., Schmidt, A.U., Shah, Y.: Smart OpenID: a smart card-based OpenID protocol. In: Gritzalis, D., Furnell, S., Theoharidou, M. (eds.) SEC 2012. IFIP AICT, vol. 376, pp. 75–86. Springer, Heidelberg (2012)
14. Morgan, R.L., Cantor, S., Carmody, S., Hoehn, W., Klingenstein, K.: Federated security: The shibboleth approach. EDUCAUSE Quarterly 27(4) (2004)
15. Pilar, D.R., Jaeger, A., Gomes, C.F.A., Stein, L.M.: Passwords usage and human memory limitations: A survey across age and educational background. PLoS One 7(12), e51067 (2012)
16. Vossaert, J., Lapon, J., De Decker, B., Naessens, V.: User-centric identity management using trusted modules. In: Camenisch, J., Lambrinoudakis, C. (eds.) EuroPKI 2010. LNCS, vol. 6711, pp. 155–170. Springer, Heidelberg (2011)
17. Yan, J., Blackwell, A., Anderson, R., Grant, A.: Password memorability and security: Empirical results. IEEE Security and Privacy 2(5), 25–31 (2004)

ErasmusApp: A Location-Based Collaborative System for Erasmus Students

Karel Bruyneel and Benedita Malheiro

Abstract. This paper reports on the design and development of an Android-based context-aware system to support Erasmus students during their mobility in Porto. It enables: (*i*) guest users to create, rate and store personal points of interest (POI) in a private, local on board database; and (*ii*) authenticated users to upload and share POI as well as get and rate recommended POI from the shared central database. The system is a distributed client / server application. The server interacts with a central database that maintains the user profiles and the shared POI organized by category and rating. The Android GUI application works both as a standalone application and as a client module. In standalone mode, guest users have access to generic info, a map-based interface and a local database to store and retrieve personal POI. Upon successful authentication, users can, additionally, share POI as well as get and rate recommendations sorted by category, rating and distance-to-user.

1 Introduction

The main goal of this application is to support Erasmus students both upon arrival and throughout their stay in Porto. Initially, newcomers can feel isolated and lost, experience language difficulties or may need to find a place to stay, i.e., they require information about where to meet young people, accommodation, food or health treatments. In a later stage, when they develop a network of relationships and establish a set of preferred indoor and outdoor places, they are ready to share this information and, thus, make recommendations to fellow students.

The developed application intends to help overcoming the initial difficulties and promoting information sharing and student integration. The ErasmusApp has

Karel Bruyneel
Katholieke Hogeschool Sint-Lieven, Ghent, Belgium
e-mail: karel.bruyneel@gmail.com

Benedita Malheiro
School of Engineering, Polytechnic Institute of Porto and INESC TEC –
INESC Technology and Science (formerly INESC Porto), Porto, Portugal
e-mail: mbm@isep.ipp.pt

L. De Strycker (ed.), *ECUMICT 2014*,
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_4, © Springer International Publishing Switzerland 2014

two operation modes: standalone and distributed. The standalone mode, which is offered to guest users or offline registered users, provides a map-based interface to explore, find and store personal POI as well as basic information regarding accommodation, restaurants, hospitals, meeting places, the local Erasmus Students Network (ESN), etc. The distributed mode requires user authentication and offers, additionally, the possibility to classify, rate and share personal POI and to get recommendations per category ordered by rating and/or distance-to-user.

This paper is organized in six sections covering the introduction, the state of the art, the development environment, the ErasmusApp system, the tests and results, the conclusions and references.

2 Context-aware Mobile Applications

There are numerous context-aware mobile applications to help tourists finding historic sites, museums, movies, shows, restaurants, stores, recreational activities, etc. PC World magazine presents in [1] a distilled list of the essential Android applications for travellers. Those that inspire this work are hereby listed:

- **COMPASS** or COntext-aware Mobile Personal ASSistant is an application that provides a tourist with information and services based on his specific context and current goal [2].
- **MyMytilene** is a mobile tourist guide platform for the Municipality of Mytilene, Greece. MyMytilene addresses personalization in the context of allowing users to explicitly select touristic content to be included in a customized mobile application which is generated on the fly, adapting the application so as to meet the screen size and hardware constraints of the user's mobile phone. The mobile application may be used in either offline or online modes [3].
- **Wikitude** is an augmented reality global travel guide that overlays Wikipedia and user-contributed content over the mobile device camera view, providing information on the user surroundings. Wikitude uses the on-board GPS, compass and movement sensors to match the user position in relation to the landmarks that the camera is pointing towards [4].
- **Hotels Near Me** uses the on-board GPS sensor to establish the user location and, then, consults a 60 000 records hotel database to find accommodation in the proximity. The user gets the ratings, address, phone number and user reviews for each hotel. Once he/she picks a hotel, the app shows a gallery of photos of the rooms and provides the price quotes for the duration of the stay. If the user actually books an accommodation, he/she receives a confirmation by e-mail. Hotels Near Me also lets the user browse for hotels prior to the trip by letting the user select 'elsewhere' from the main screen and then specifying the desired city [5].

- **Where To?** provides in real-time information on cheap gas, movies, shows, restaurants, traffic conditions, weather forecasts and news headlines [6] based on the user location.
- **Yelp** is a service, which includes an Android app, that provides location-based business information on top of real-time images from the phone's camera [7].
- **moreTourism** is a content-based and collaborative (hybrid) recommendation platform providing information about tourist resources depending on the user profile, location, schedule and the amount of time for visiting interest points isolated or combined in a route [8].
- **WebMD** is a location and preference-aware healthcare and fitness mobile application mostly for non-clinical use [9].
- **Cinemappy** is a location-based mobile application that provides enriched contextual movie recommendations. The current spatial and temporal user location is the basis for the contextualization and DBpedia, one of the best-known datasets publicly available in the Linked Open Data (LOD) project, is the semantic enrichment source [10].
- **Where** is a mobile application supported by a local search and recommendation portal that allows people build their own profiles and to find local events, read reviews and other business information, and write their own reviews [11].

All these context-aware mobile applications provide one or more categories of information related to the current user context. The recommendation categories range from health [9], tourism [2,4,3,5,8], business [4,6,7,11] to media content [10] and their sophistication varies from text, image, video or augmented reality suggestions. The user context information corresponds overwhelmingly to the user current spatial and temporal data.

The idea behind the Erasmus App is close to the "Where" or "Where To?" applications since it provides and allows the classification of location-based recommendations regarding user added locations or POI. However, the Erasmus App, which is the front-end of a dedicated location-based collaborative tool, does not include the remaining functionalities of these products since it focussed on creating a sense of community and a support network for Erasmus students.

3 Development Environment

The development of Android applications is a dynamic field involving several technologies and areas of expertise. Therefore, the first task of an Android developer is to select the set of inputs, languages, technologies, tools and API to use. Since the Android programming environment is Java-based, the distributed system (client and server modules) was fully developed in Java, using Eclipse as the Integrated Development Environment (IDE).

In terms of persistent data storage, the user local database relies on the local Android device SQLite database engine [12]. The central database is stored in a MySQL database server [13]. The communication with both SQLite and MySQL uses Structured Query Language (SQL) commands.

A dedicated object oriented protocol was created and used for the communication between the GUI application and the server. All exchanged messages extend from one class with undefined objects as attributes, allowing the definition of specific subclasses for the different message primitives to be exchanged.

To enable the installation of the application in as many devices as possible, the developed code only uses the Android application programming interface (API) 8 and the Google Android API for map views [14]. In particular, the Android API 8 is supported by Android devices with version 2.2.1 or higher [15].

To determine the user context, which in this case is the user location, both the GPS and the network providers are used [16]. The selection of the provider depends on whether the user is connected to the Internet or whether the user is indoors or outdoors. When displaying a map, the compass input is also used to show the North. The inputs of the accelerometer and the gyroscope have been studied and considered, but are not used in this version of the application.

4 ErasmusApp Location-Based Collaborative System

The ErasmusApp system is composed of the front-end mobile application, which includes a local database, and the back-end server module together with the central database. The system requires Internet access to be fully operational, i.e., to provide add, delete, edit, share and rate POI functionalities. Furthermore, the user can request a map view of all locations.

4.1 Architecture

Figure 1 shows the overall architecture that includes the Android GUI application together with the local database, the dedicated server module and the central database. The GUI application requires an Internet connection to interact with the dedicated server module, the Google Maps servers and all external sites used. Every module plays a well-defined role in the overall system:

- The Android application provides the graphical user interface (GUI), including, the Web and map views. The Web views correspond to the *Basic Erasmus info* Web links and the map views display the maps provided by Google Maps.
- The local device database contains a set of predefined locations necessary to provide the *Basic Erasmus info* as well as all personal user-defined POI.

- The server module processes the client requests and interacts with the database server using SQL. The client/server communication uses a dedicated object-oriented protocol.
- The central database contains all information about the registered users, shared POI, ratings, etc.

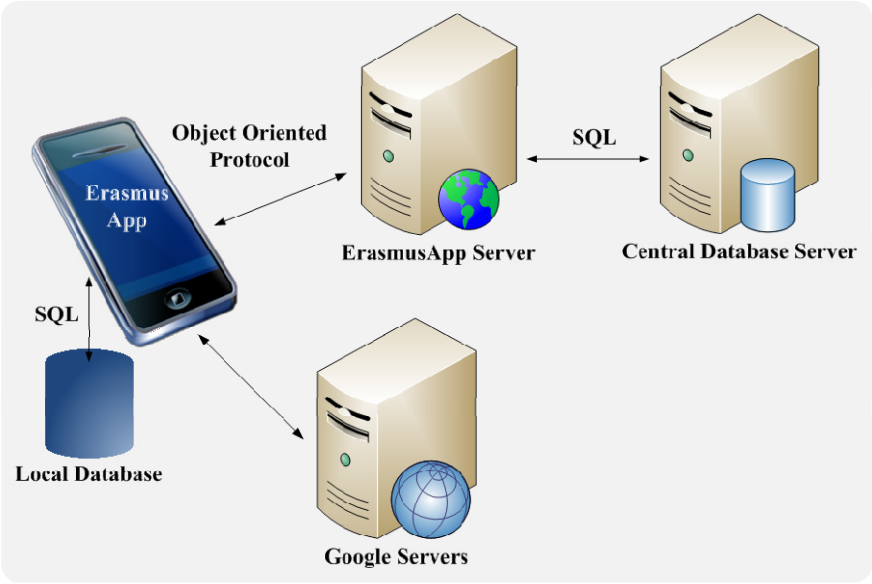


Fig. 1 ErasmusApp system architecture

4.1.1 Persistent Storage

The enhanced entity–relationship (EER) model of the MySQL central database in crow's foot notation is shown in Figure 2.

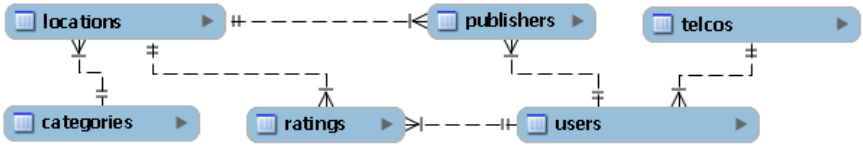


Fig. 2 Central database EER model

The main central database tables are the *locations* (POI) and the *users* tables, which are connected via the *ratings* and *publishers* tables. Additionally, every user subscribes a given telecommunications company (telco) and every location belongs to a category from the *telcos* and *categories* tables, respectively. The

publishers table allows a location to have multiple publishers (users). The *ratings* table stores the location ratings, where multiple users can rate a single POI and a user can rate multiple POI. The *categories* table holds the eight pre-defined POI categories: hospitals, supermarkets, restaurants, hairdressers, outside & quiet, touristic, nightlife and others.

The on-device SQLite database, which stores the personal POI, holds a single *locations* table with id, name, description, latitude and longitude attributes.

4.1.2 Object Oriented Protocol

The application level protocol used for the communication between client and server modules contains several messages with contents depending on their purpose. Figure 3 shows the contents of the protocol package. This package provides the client-side and server-side modules with a common understanding of the exchanged messages and objects.

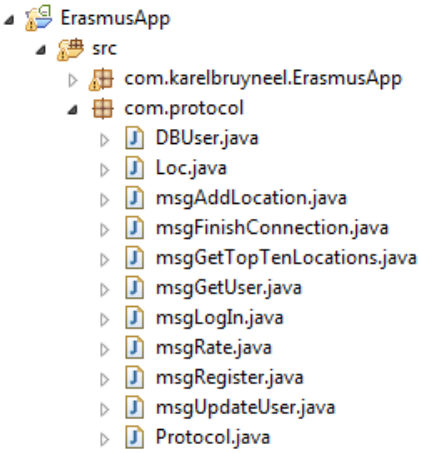


Fig. 3 Protocol package

The package defines the *DBUser* and *Loc* objects which represent a user and a location (the object *Location* is already defined in the Google Android API). Furthermore, it specifies a class called *Protocol*, which is inherited by all protocol messages, with four generic objects that will be latter casted to specific object types in the message subclasses.

After messages are exchanged, the connection between client and server closes in order to minimize the cost of communication.

4.2 Distance-to-User

The equirectangular projection distance (in meters) between the user current position and a given POI is calculated using Equation (1) and Equation (2):

$$\phi_m = (lat_u + lat_{POI})/2 \tag{1}$$

$$d = R\sqrt{(lon_u \cos\phi_m - lon_{POI} \cos\phi_m)^2 + (lat_u - lat_{POI})^2} \tag{2}$$

where ϕ_m is the mean latitude angle, R is de radius of the Earth (in meters) and lon_u, lat_u, lon_{POI} and lat_{POI} represent the geodetic longitude and latitude coordinates of the user *u* and *POI*.

When the user requests recommendations, they are selected and ordered by rating (average number of stars) and distance-to-user (*d*). Therefore, the application recommends POI ordered by the quotient between the average rating and the distance-to-user (stars/m).

4.3 Functionalities

The ErasmusApp has a dual-mode operation: standalone (guest users) or client (registered users) application. In standalone mode, guest users have access to generic info, a map-based interface and to the local database to store and retrieve personal POI. In distributed mode, authenticated users can, additionally, share POI as well as get and rate recommendations sorted by category, rating and distance-to-user. The stand-alone mode is intended to minimize connection costs and for offline operation. The GUI application offers the following functionalities:

- In stand-alone mode to guest users:
  - *Map* shows the user location on a Google map.
  - *Basic Erasmus info* loads an initial set of POI, including essential locations such as the *Closest Hospitals*, Porto *ESN* and *Looking for a House?* information, as well as famous student meeting points like *Piolho*, a downtown coffee house, *Cais da Ribeira*, the old quay by the river, and *Matosinhos beach*. This option provides, apart from the regular Android activities pages, a Web view activity.
  - *Personal Locations* allows the user to: (*i*) add, edit and delete personal POI to and from the local database and (*ii*) display personal POI on a map.
- In distributed mode and in addition:
  - *Login* for user authentication and a registration button.
  - *Create/Edit profile* allows the user to register and modify his profile.
  - *Find Location* recommends the top ten shared POI per category. These recommendations can be ranked by distance-to-user, rating or both rating

and distance-to-user. The user, after receiving the top ten recommended locations, can go to the POI page to rate and view the location on a map.

– *Personal Locations* allows the user to: (*i*) add, edit and delete personal POI to the local database; (*ii*) display personal POI on a map; and (*iii*) categorise, rate and share the POI with fellow users (only rated categorised POI can be shared). When adding a new POI, its position can be determined automatically via the GPS sensor, the network provider or established by the user by clicking on the map.

5 Tests and Results

Tests were successfully conducted both with the Android device and the Eclipse emulator plug-in. Although the application runs faster on the device, the emulator was used to capture higher quality screenshots.

Map

The mapping functionality is widely used to display the user and POI over Google Maps. Figure 5 displays a tapping map of the user location and other POI.

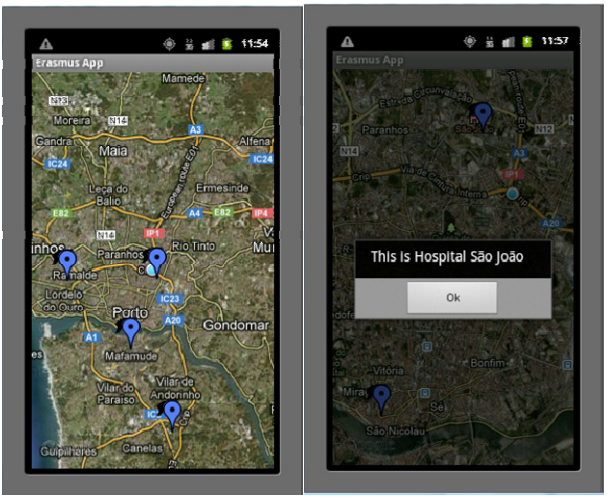


Fig. 4 Displaying the user location with a tapping map

Basic Erasmus info

The *Basic Erasmus info* loads a default set of POI, including hospitals, accommodation, Erasmus-related information and student meeting points. Figure 6 displays the ESN info, ESN site and a map with the user and the closest hospitals.

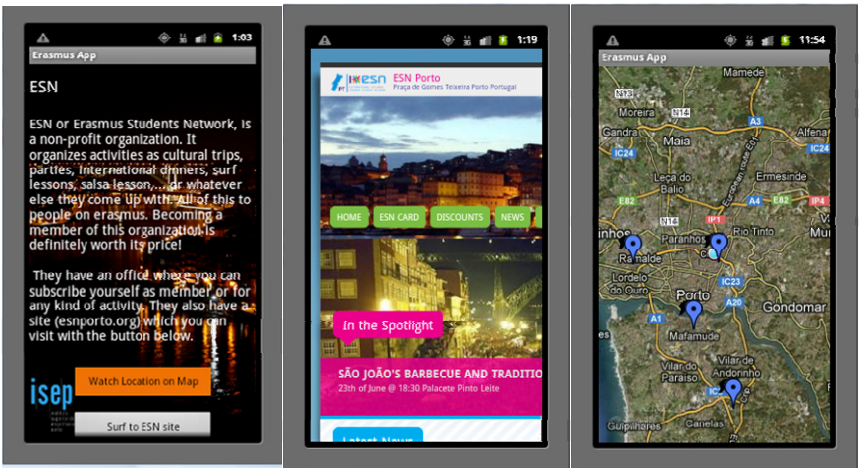


Fig. 5 Displaying basic information

Registration and Profile Editing

User registration and profile editing share the same screen as shown in Figure 7.

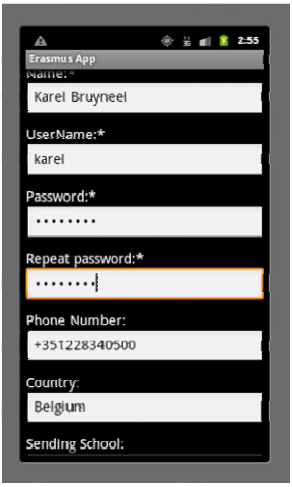


Fig. 6 Registration/Profile editing

Personal Locations

Personal POI can be added using the current user location, specifying the geodetic coordinates or via a clickable map – Figure 8. When the user decides to share a personal POI, first has to attribute a category and a rating to the POI and, then, has to upload the POI to the central database – Figure 8. Upon success, the user gets a new page displaying the new shared POI together with the attributed rating.

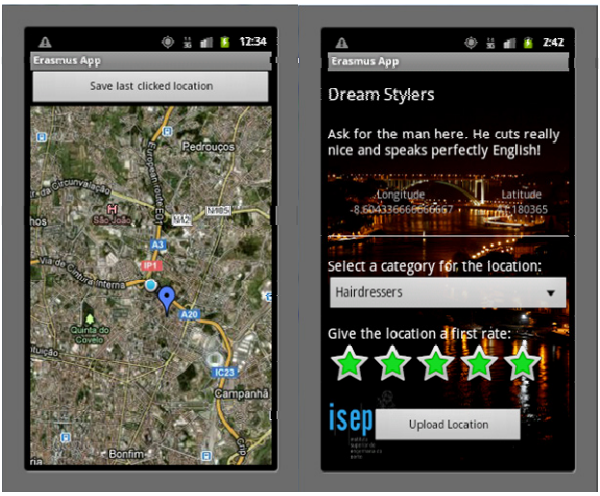


Fig. 7 Adding a personal POI via a clickable map and sharing a personal POI

Find Locations

Find locations returns the ten best ranked shared locations per category. The ranking is based on the collaborative rating and/or on the distance-to-user. Figure 9 shows the result of a *Find locations* request for supermarkets ordered by rating together with a detailed view of the highest ranked supermarket. The detailed view presents the name, description, coordinates and the collaborative rating of the item. Additionally, the user can add/edit his personal rating, which is immediately incorporated into the collaborative rating. The button on the bottom of the page loads a map displaying this POI and the user location.

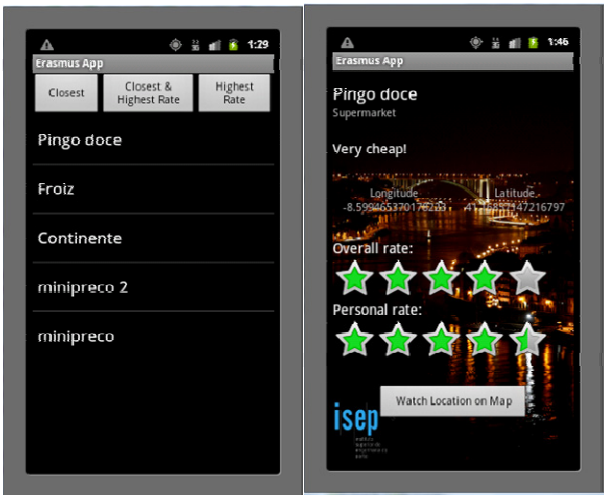


Fig. 8 Recommending shared supermarkets by rating

6 Conclusion

The application is intended to foster the integration of Erasmus students by collaboratively creating and rating a POI database for current and future usage. Users decide whether to keep or share their personal POI.

Achievements

The ErasmusApp system is a distributed application composed of a central database, a server module and a mobile map-based GUI application with a local database. The Android-based GUI application allows both standalone and client operation modes. Guest users can add, rate, edit, display on a map and delete personal POI that are stored in the local database as well as register. Authenticated users can edit the profile, share personal POI, by uploading them to the central database, and get recommendations by category ranked by rating, distance-to-user or the combination of both. Shared POI are not editable nor removable by regular users.

The communication between client and server and between server and database are both supported by TCP. While the first uses a dedicated protocol, the latter exchanges SQL commands. The user spatial context is established through the GPS and/or the network provider.

The tests conducted showed that the implemented functionalities work properly. The application can be easily ported for other cities than Porto since only the basic info needs to be substituted.

Future Work

The system has several limitations and new functionalities are currently being developed, including geo-referenced video sharing, semantic enrichment via LOD repositories and hybrid recommendations, i.e., that take into account both the user personal profile and the fellow users ratings. In terms of personal profile, it is being refined to accommodate both spatial and temporal context data, e.g., use the temporal context to make recommendations based on the period of the day, day of the week or season, and the past POI ratings by category. Another promising feature is to add a new accommodation main entry in the start menu, rather than in the basic Erasmus info section, and expand it to allow finding flats and flatmates.

More work has to be done on reporting invalid or duplicate locations. The server should look not only for identical names, which it does right now, but, also, for nearby identical category POI. In this case, the user should get a list of identical close POI for selection before confirming his own POI input. Additionally, the user should be able to report incorrect data, e.g., wrong addresses. This report would then notify the publisher of the fact and grant the publisher the permission to correct his input. At a later stage, the application could compute the credibility of the publishers and, thus, use it to compute the POI rating.

Finally, user assessment and feedback has to be performed to refine and improve the ErasmusApp.

Acknowledgments. This work was partially supported by the ERDF – European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project «FCOMP – 01-0124-FEDER-022701». The authors would like to thank their home institutions, ISEP and KAHOSL, and the EU Lifelong Learning Programme, in particular, the Erasmus student mobility programme.

References

1. PCWorld, 11 Essential Android Travel Apps (2010), http://www.pcworld.com/article/193523/11_essential_android_travel_apps.html (accessed in June 2012)
2. van Setten, M., Pokraev, S., Koolwaaij, J.: Context-aware recommendations in the mobile tourist application COMPASS. In: De Bra, P.M.E., Nejdl, W. (eds.) AH 2004. LNCS, vol. 3137, pp. 235–244. Springer, Heidelberg (2004)
3. Gavalas, D., Kenteris, M.: A web-based pervasive recommendation system for mobile tourist guides. Personal Ubiquitous Computing 15(7), 759–770 (2011), doi:10.1007/s00779-011-0389-x, ISSN 1617-4909
4. Wikitude, Wikitude - The World's leading Augmented Reality SDK (2012), http://www.wikitude.com/ (accessed in June 2012)
5. Bluemedialab.com, "Hotels Near Me" App. (2012), http://blumedialab.com/products/google-android-apps/Hotels-near-me-for-Android.htm (accessed in June 2012)
6. FutureTap GmbH, "Where To?" App. (2012), http://www.futuretap.com/apps/whereto-en/ (accessed in June 2012)
7. Yelp Inc., Yelp. (2012), http://www.yelp.com/ (accessed in June 2012)
8. Rey-López, M., Barragáns-Martínez, A.B., Peleteiro, A., Mikic-Fonte, F.A., Burguillo, J.C.: moreTourism: Mobile recommendations for tourism. In: 2011 IEEE International Conference on Consumer Electronics (ICCE), pp. 347–348 (2011), doi:10.1109/ICCE.2011.5722620, ISSN 2158-3994
9. Liu, C., Zhu, Q., Holroyd, K.A., Seng, E.K.: Status and trends of mobile-health applications for iOS devices: A developer's perspective. Journal of Systems and Software 84(11), 2022–2033 (2011), doi:10.1016/j.jss.2011.06.049, ISSN 0164-1212
10. Ostuni, V.C., Di Noia, T., Mirizzi, R., Romito, D., Di Sciascio, E.: Cinemappy: a Context-aware Mobile App for Movie Recommendations boosted by DBpedia. In: Proceedings of the International Workshop on Semantic Technologies Meet Recommender Systems & Big Data (SeRSy 2012), pp. 37–48 (2012), http://ceur-ws.org/Vol-919/paper4.pdf ISSN 1613-0073 (accessed in December 2012)
11. Where, "Where for Android" (2012), http://site.where.com/download-where/google-android/ (accessed in June 2012)
12. Vogel, L.: Android SQLite Database and Content Provider - Tutorial (2012), http://www.vogella.com/articles/AndroidSQLite/article.html (accessed in June 2012)
13. MySQL, The world's most popular open source database (2012), http://www.mysql.com/ (accessed in June 2012)

14. Volgel, L.: Location API and Google Map in Android (2012), http://www.vogella.com/articles/AndroidLocationAPI/article html (accessed in June 2012)
15. Android Developers, "Platform Versions" (2012), http://developer.android.com/about/dashboards/index.html (accessed in June 2012)
16. Android Developers, "Location Strategies" (2012), http://developer.android.com/guide/topics/location/strategies.html (accessed in June 2012)

Smart Object for 3D Interaction

Hannes Harms, Toomas Juht, Anna Janaszkiewicz, Jana Valauskaité, António Silva, Benedita Malheiro, Cristina Ribeiro, Manuel Silva, Nídia Caetano, Paulo Ferreira, and Pedro Guedes

Abstract. This paper reports on the creation of an interface for 3D virtual environments, computer-aided design applications or computer games. Standard computer interfaces are bound to 2D surfaces, e.g., computer mouses, keyboards, touch pads or touch screens. The Smart Object is intended to provide the user with a 3D interface by using sensors that register movement (inertial measurement unit), touch (touch screen) and voice (microphone). The design and development process as well as the tests and results are presented in this paper. The Smart Object was developed by a team of four third-year engineering students from diverse scientific backgrounds and nationalities during one semester.

1 Introduction

Progress in technology, especially in fields of embedded systems, communication and sensor systems, give rise to new types of applications and devices. The availability of low cost Micro-Electrical-Mechanical Systems (MEMS), which are very small electro-mechanical devices, allows on chip physical measurements such as acceleration or rotation. This has enlarged the scope of application domains and fostered the engineers ability to combine several sensor technologies to test new methods of human computer interactions.

This project was focused on the design and development of a Smart Object to be used as a 3D Personal Computer (PC) interface. Standard computer interfaces are bound to 2D surfaces, e.g., the computer mouse which gives allows movement in the X and Y axis. The goal of this project is to design a product that allows 3D movement and, moreover, that enables the user to turn, rotate, zoom in and out on a computer screen. This type of device can be useful in computer games with 3D

Hannes Harms · Toomas Juht · Anna Janaszkiewicz · Jana Valauskaité · António Silva · Benedita Malheiro · Cristina Ribeiro · Manuel Silva · Nídia Caetano · Paulo Ferreira · Pedro Guedes
European Project Semester, School of Engineering, Polytechnic Institute of Porto, Porto, Portugal
http://www.eps2012-wiki3.dee.isep.ipp.pt
e-mail: epsatisep@gmail.com

L. De Strycker (ed.), *ECUMICT 2014,*
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_5, © Springer International Publishing Switzerland 2014

graphics, on which often there is the need to simultaneously translate and rotate the virtual world, but can also be of extremely utility in professional Computer Aided Design / Computer Aided Manufacturing / Computer Aided Engineering (CAD / CAM / CAE) applications. Very often, the users of these professional applications need to translate, rotate, zoom in and out the view of the products that they design in CAD applications (such as, just as an example, SolidWorks™ [1]) and the virtual equipment's to manufacture these products, that are modelled virtually in applications such as ABB RobotStudio (an application for robot simulation and off line programming from ABB [2]) or, even, computer applications for layout design (such as QUEST, from DELMIA [3]). The fast and simultaneous translation, rotation and zooming of the virtual worlds in these types of applications is often difficult to realize and even gives rise to some difficulties that could be surpassed by using and Human-Computer interface such as the Smart Object herein described.

This paper is structured in five sections. The introduction (Section 1) presents the general problem, motivation and objectives and is followed by the description of related three-dimensional PC interfaces (Section 2). Section 3 introduces the support technologies adopted. Section 4 addresses the design and construction of the device, the control software, the data connection and the representation of the object in a virtual 3D environment. Finally, Section 5 discusses the results and presents the conclusions.

2 3D Interactive Objects

The 3D mouse from 3Dconnexions is a family of devices with identical functionalities. The company 3Dconnexions, which has launched the first product in 2009, has five different products that can be used for the same purpose, but with diverse segmentation [4]. They differ in size and functionalities, e.g., the products with higher dimensions have a built-in liquid crystal display (LCD), have more functions and are more expensive. The smallest device has just a Controller Cap and two interface buttons. The 3D mouse device from 3Dconnexions, that is identical to the Smart Object, resembles a joystick and allows the user to rotate, zoom, turn and pan the object. The manufacturer describes this 3D mouse as having "a level of control that is simply not possible with a traditional mouse and keyboard" [4].

A cube, equipped with an Arduino based development board, various sensors and an XBee wireless interface was developed by [5]. This cube was tested as an input controller for several computer games and the software was released under an open source license.

3 Support Technologies

The technologies that support this project are standards in modern electronic devices and are mass produced.

3.1 Microcontroller

The Arduino is a popular open source single-board microcontroller designed to make the process of using electronics in multidisciplinary projects more accessible [6]. The hardware consists of simple open hardware design for the Arduino board with an Atmel AVR processor (it is a single chip microcontroller which was developed by Atmel in 1996) and on-board input/output support [7]. The software consists of a standard programming language compiler and the boot loader that runs on the board.

3.2 Sensors

An inertial measurement unit (IMU) is an electronic device that measures and reports the velocity, orientation and gravitational forces applied to a craft and uses a combination of accelerometers and gyroscopes. Typically, they are used to maneuver aircraft, including unmanned aerial vehicles (UAV), and spacecraft, including shuttles, satellites and landers. The IMU is also the main component of the navigation systems used in aircrafts, spacecrafts, watercrafts and guided missiles. In this capacity, the data collected from the IMU sensors allows tracking a craft. This project uses an IMU to sense movement.

The touch screen is an electronic visual display that detects the presence and location of a touch within the display area. Although the term generally refers to touching the display of the device with a finger or hand, it can also sense other passive objects such as a stylus. In this project the touch screen is intended to detect the touch of the user's finger.

A microphone is an acoustic-to-electric transducer or a sensor that converts sound into an electrical signal. In this project the microphone is intended to detect and convert the user voice into an electrical signal.

3.3 Power Supply

To power such a device it is possible to use solar panels or rechargeable batteries since the remaining options are too difficult, expensive or complex to implement.

Solar panels are a packaged, connected assembly of photovoltaic cells. The solar panel can be used as a component of a larger photovoltaic system to generate and supply electricity for commercial and residential applications. A photovoltaic system typically includes an array of solar panels, an inverter and, sometimes, a battery and/or solar tracker. Solar panels are expensive when compared with batteries. Besides, since the Smart Object is intended to be held by the user, it should not be too large. Additionally, the solar panels would often be obstructed by the user hands and, in contact with human skin, would require frequent cleaning.

Lithium Polymer (LiPo) batteries are rechargeable batteries composed of several identical secondary cells in parallel to increase the discharge current capability. They are compact and ideal for small cordless devices. This project uses a LiPo battery as power source.

3.4 Wireless Data Connection

Wi-Fi is a technology that allows an electronic device to exchange data wirelessly over a data network, including high-speed Internet connection. There are a lot of devices that can use Wi-Fi, such as PC, video game consoles, smartphones, tablets or digital audio players. All of these can connect to a network resource such as the Internet via a wireless network access point. Indoors the range is about 20 m and outdoors is greater. Access point coverage, also called hotspots, can comprise an area as small as a single room with walls that block radio waves or as large as many square miles - this is achieved by using multiple overlapping access points [8].

ZigBee is intended to be simpler and less expensive than other wireless personal area network (WPAN) such as Bluetooth [9]. It is targeted at radio-frequency applications that require low data rate, long battery life and secure networking. ZigBee is best suited for periodic or intermittent data or single transmission from a single transmission from a sensor or input device [10].

Bluetooth is a widely spread wireless connection option [11]. It can be found on PC, smartphones, laptops and other devices. It is used for exchanging data over short distances from both fixed and mobile devices. Bluetooth creates personal area networks (PAN) with high levels of security [12].

Bluetooth was selected for the wireless data link because it is present in most laptops, tablets and smartphones platforms. ZigBee was contemplated as an alternative wireless technology, but was not adopted because requires acquiring a receiver for the user platform. Finally, Wi-Fi was discarded since it is, among the wireless technologies, the biggest consumer of energy.

4 Project Development

The system is made of the Smart Object, the wireless data link and the PC application. On the Smart Object side, the Arduino microcontroller, that is connected to the sensing system, is responsible for collecting and transmitting all measurement data over the wireless data link to the PC application. The sensing system includes an IMU, that provides a full six degree of freedom measurement of the position of the object, an accelerometer, a microphone and a touch pad. On the PC-side, the object is modelled and displayed in a virtual environment using Java programming language.

4.1 System Components

All components were selected considering price, ease of use, flexibility and energy consumption. The Smart Object includes:

- **Arduino BT** – a microcontroller board based on the ATmega328 and the Bluegiga WT11 Bluetooth module. Its main features are the wireless serial communication over Bluetooth, 14 digital input/output pins, 6 analog inputs, a 16 MHz crystal oscillator, a 32 kB Flash Memory and 2 kB SRAM [13].
- **IMU Digital Combo Board ITG3200/ADXL345** – a six degrees of freedom measurement with an I^2C interface containing:
  - ITG-3200 Integrated Triple-Axis Digital-Output Gyroscope is a single-chip, digital-output, 3-axis MEMS gyro optimized for gaming, 3D mice and motion-based remote control applications. It features three 16-bit analog-to-digital converters (ADC) for digitizing the gyro outputs and a Fast-Mode I^2C (400 kHz) interface. The Digital-output *X*, *Y* and *Z* axis angular rate sensors have a sensitivity of 14 LSB/°/s and a full-scale range of $\pm$2000°/s [14].
  - ADXL345 Accelerometer – a small and ultra low power 3-axis accelerometer. It measures the static acceleration of gravity as well as the dynamic acceleration resulting from motion or shock. The measurement is performed with a high resolution (13-bit) at up to $\pm$16 g. The data is formatted as 16-bit two's complement and is accessible through a I^2C digital interface. The resolution of 3.9 mg/LSB enables measurement of inclination changes less than 1.0°[15].
- **ADMP401 MEMS Microphone** – an analog omnidirectional microphone [16].
- **Nintendo DS Touch Screen LCD-08977** – a 4-wire analog touch screen manufactured by Hantouch USA originally for the Nintendo DS [17].

4.2 Data Flow

The Smart Object detects movement, sound and touch sensing and communicates all perceived changes to the PC application via the wireless data link – Fig. 1.

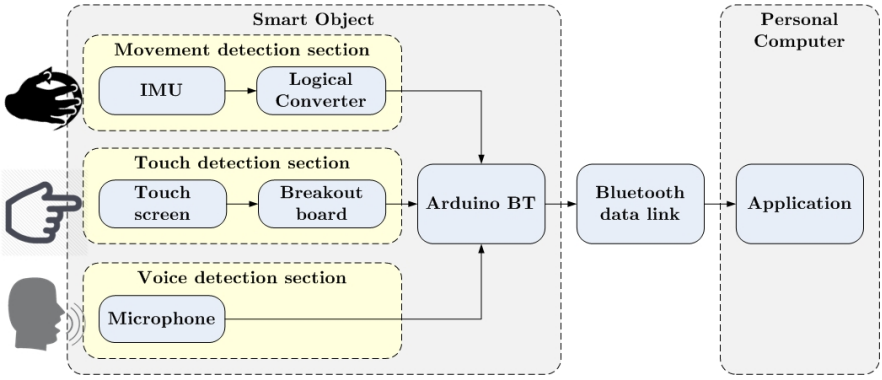


Fig. 1 Data flow

The IMU, microphone and touch pad are continuously measuring and sending the detected user inputs to the Arduino: (*i*) the IMU senses and reports the movement-related data to the Arduino via a logical level converter; (*ii*) the microphone registers and sends directly the voice input to the Arduino; and (*iii*) the touch pad detects the user input (touch) and sends it via a breakout board to the Arduino.

4.3 Software Modules

The Smart Object programming was divided in two sections. On the device side, the Arduino program was written in C and, on the PC side, was coded in Java. To get powerful graphical functions the PC-side application uses a Java OpenGL engine called "jmonkeyengine" [18]. Eclipse was use for the Java programming development. Figure 2 gives an overview of the software architecture. The Java application is composed of three threads: (*i*) the data management thread for the Bluetooth data link; (*ii*) the sensor data plotting thread; and (*iii*) the 3D-environment visualization thread.

4.3.1 Arduino Application

The Arduino application is structured in a setup and a loop functions. In the setup, the serial connection is established, all digital and analog ports, which are connected to the sensors, are defined and an array of double values to hold the sensor values

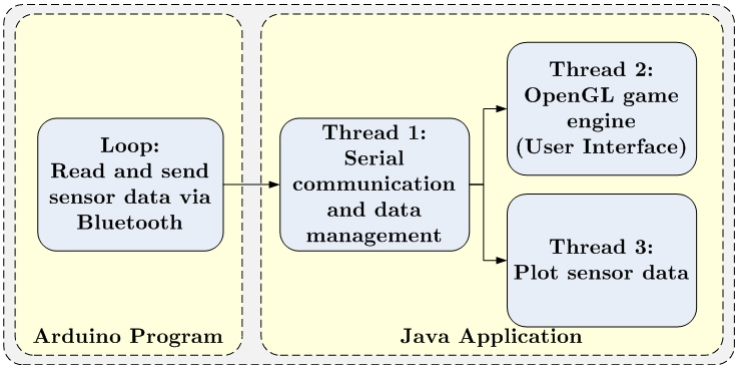


Fig. 2 Software architecture

is declared. After the execution of the setup, the infinite control loop starts. If a wireless connection with the PC exists, the Arduino reads all sensor data. In the case of the IMU, an open source library called "FreeIMU" [19] was used. The gyro values and acceleration values are fused together to calculate a more stable position. As a result, it is possible to obtain directly the Euler Angles from the IMU. The touch pad is directly connected to an analog port and requires a special pin configuration to allow reading the X and Y position simultaneously. Finally, the microphone analog output is read.

4.3.2 PC Application

The PC application is responsible for establishing the wireless data link (Data Management Thread), plotting all measurement data (Plot Thread) and, finally, displaying the Smart Object in the 3D virtual environment (Game Engine Thread).

Data Management Thread

The data management thread establishes the connection between the PC application and the Smart Object. The Bluetooth profile used is a serial connection. If the Bluetooth connection is established, all incoming data is stored in three different First In First Out (FIFO) queues. These FIFO are declared as public. The implemented data structure is thread safe in order to avoid data crashes since multiple threads access the incoming data.

Plot Thread

The plot thread plots incoming sensor data in real time. The thread associates a time stamp to each incoming value as soon as it is added to the data management thread FIFO and refreshes the corresponding plot. All plots are accessible via a tabbed pane

in the Java application, allowing the user to switch between the different sensor data plots and the 3D environment - Fig. 3.

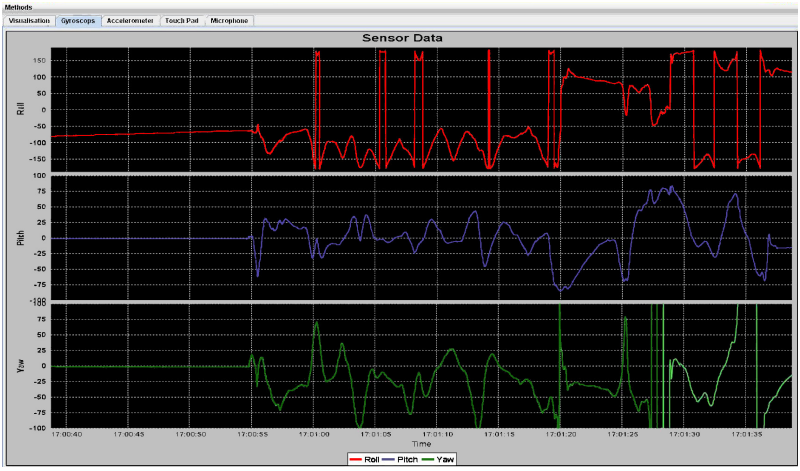


Fig. 3 Plot thread visualisations

Game Engine Thread

The Game Engine Thread implements the OpenGL game engine. After setting up the 3D virtual environment and displaying the cube, the thread tries to get data from the data management thread. If data is available, a loop is started to permanently update the visualization. Figure 4 shows the flowchart of this thread. To detect the zoom, the touch pad data is analysed in a separate function. If the values are decreasing or increasing the zoom value is calculated accordingly. By analysing the acceleration data for big changes in a short time period, a shake of the object can be detected. With this information two modes are implemented. The first mode rotates the cube and the second mode rotates the camera around the cube. For both modes the three Euler angles from the IMU are used to change the rotation of the cube or camera.

Figure 5 shows the 3D virtual representation of the Smart Object.

4.4 Building and Assembly

Finally, the project development involved building the external cubical container and the assembly of the components.

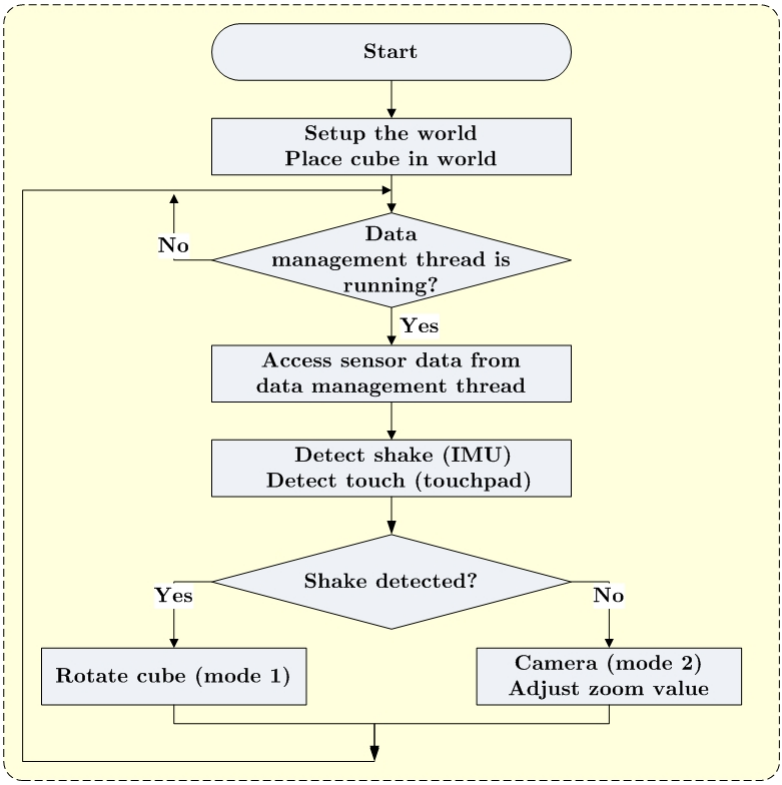


Fig. 4 Game engine thread

Fig. 5 3D Virtual environment

4.4.1 Angle Brackets

All Smart Object components, except the touch screen, are stored inside the container. The outer shell is made from acrylic glass that was machined and bent to the needed shape. The angle brackets were made from an aluminium alloy L-profile. The L-profile had to be cut in 20 mm wide pieces and then two holes had to be drilled into each angle bracket for the screws. The first hole drilled was 2 mm in diameter and the second one was in 3 mm in diameter. Rasps were used to make the edges surfaces smoother, because a metal saw was used for the cutting. Angle brackets for the bottom part had to be smaller from the 2 mm hole side so that the Arduino could be taken out if needed. Thus, the angle brackets were grinded to size.

4.4.2 Bending the Acrylic

Before bending the acrylic, the container parts were cut to size and small incisions of 0.3 mm to 0.5 mm deep were made along the bending lines. These incisions were performed on the inner side of the bent in order to relieve stress that the bending causes. A voltage regulator was used to heat up a wire that was fastened between two clamps. The wire was placed 3 mm from the piece of acrylic sheet and was heated to approximately 400 °C (the voltage regulator was adjusted to 3.1 V and 11.8 A). This temperature was best suited for the 3 mm thick acrylic glass sheet used.

4.4.3 Machining the Acrylic

The holes for the screws were drilled after bending the acrylic parts. Chamfers were made so that the screw heads would be at the same level as the acrylic. A drilling machine was used for this purpose. Additional holes were made for the touch screen connector and for the USB cable to charge the assembled device. The top and bottom parts were cut into size, the sides and corners were grinded and filed. Finally, the holes for the screws were drilled. Chamfers were also made to hide the screw heads.

4.4.4 Assembly

To assemble all parts different screws, washers and nuts were used. The angle brackets hold together 4 acrylic glass parts that form the shell. Two types of screws are used to fasten the angle brackets to the acrylic: metal screws and M3 screws. The metal screw is screwed straightly in the angle bracket's 2 mm diameter hole. The M3 screw goes through the 3 mm diameter hole and is fastened with a nut. To fasten all the electronic things to the acrylic M3 screws, rubber washers and M3 nuts are used. The M3 screw is fastened to the acrylic with a nut. This method should relieve some stress from the electronic parts. Then the electronic parts are fixed between 2

rubber washers and to hold them in place another nut is used. The last nut is loosely fastened on the screw to keep the hardware in place. The touch screen is taped onto one side of the shell. The touch screen pin goes through the hole that was made before into the acrylic glass. The connector is placed inside the shell. Figure 6 shows the Smart Object with all assembled parts.

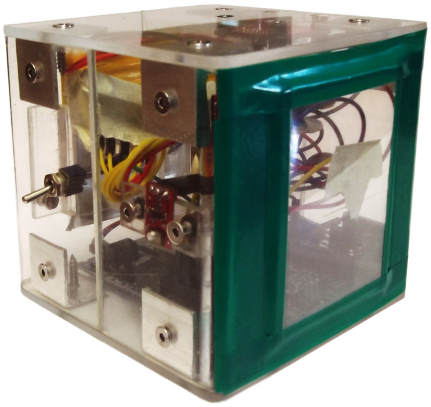


Fig. 6 The Smart Object with all assembled parts

5 Conclusion

This paper presents the Smart Object developed to interface with a 3D virtual environment. The system consists of three main modules: (*i*) an Arduino microcontroller connected to a set of sensors including an IMU that gets six degrees of freedom measurement of the position of the object; (*ii*) a Bluetooth wireless data connection; and (*iii*) a 3D virtual environment which was programmed in Java. The latter is to be installed in a laptop, PC or tablet with a Bluetooth interface.

The prototype needs, in order to become a product, to be redesigned in size, aesthetic and electronics terms. At the moment the prototype shell is made of acrylic glass, which is not suitable for mass production. The shell should be made from a moldable material like acrylonitrile butadiene styrene (ABS) plastic, which is a robust material with good impact resistance that can be easily molded with proper machines. As far as the electronic components are concerned, although the Arduino BT is ideal for prototyping, the design of a dedicated circuit board including all electronic components is necessary to reduce the size, weight and cost. The battery should also be more compact.

Acknowledgements. The authors would like to thank their home institutions and the EU Lifelong Learning Programme, in particular, the Erasmus student mobility programme.

References

1. 3D CAD Designing Software Systems, http://www.solidworks.com/ (accessed: January, 2014)
2. RobotStudio - Robotics, http://new.abb.com/products/robotics/robotstudio (accessed: January, 2014)
3. Digital Manufacturing & Production - DELMIA - Dassault Systèmes, http://www.3ds.com/products-services/delmia/ (accessed: January 2014)
4. 3D connexion, What is 3D mouse, http://www.3dconnexion.com/products/what-is-a-3d-mouse.html (accessed: April, 2012)
5. Doggen, J., Neefs, J., Brands, E., Peeters, T., Bracke, J., Smets, M., Van der Schueren, F.: Smart Objects for Human Computer Interaction, Experimental Study. In: de Strycker, L. (ed.) Proceedings of the Fifth European Conference on the Use of Modern Information and Communication Technologies (ECUMICT 2012), pp. 117–128. Ghent, Belgium (2012) ISBN 978-9-08-082554-3
6. Margolis, M.: Arduino Cookbook. Oreilly & Associates Inc., Sebastopol (2012) ISBN 978-1-449-31387-6
7. Wheat, D.: Arduino Internals. Apress, New York (2011) ISBN 978-1-4302-3882-9
8. Faludi, R.: Building Wireless Sensor Networks. O'Reilly Media, Sebastopol (2010) ISBN 978-0-596-80773-3
9. Gislason, D.: ZigBee Wireless Networking (2004) ISBN 978-0-7506-8597-9
10. Farahani, S.: ZigBee Wireless Networking and Tranceivers. Newnes, Burlington (2008) ISBN 978-0-7506-8393-7
11. Kammer, D., McNutt, G., Senese, B.: Bluetooth Application Developer's Guide: The Short Range Interconnect Solution. In: Bray, J. (ed.). Syngress, Rockland (2002) ISBN 1-928994-42-3
12. Huang, A.S., Rudolph, L.: Bluetooth Essentails for Programmers. Cambridge University Press, New York (2007) ISBN 978-0-521-70375-8
13. Arduino, A.B.T.: (accessed: April, 2012), http://arduino.cc/it/Main/ArduinoBoardBluetooth
14. ITG-3200 Integrated Triple-Axis Digital-Output Gyroscope, http://www.invensense.com/mems/gyro/itg3200.html (accessed: April, 2012)
15. Analog Devices, ADXL345: 3-AXIS Digital accelerometer, http://www.analog.com/en/mems-sensors/mems-inertial-sensors/adxl345/products/product.html (accessed: June, 2012)
16. Analog Devices, ADMP401: Omnidirectional Microphone with Bottom Port and Analog Output, http://www.analog.com/en/mems-sensors/mems-microphones/admp401/products/product.html (accessed: June, 2012)

17. Hantouch USA, How it works: 4-wire Analog-Resistive Touch Screens, http://tronixstuff.com/wp-content/uploads/2010/12/hantouch-data-sheet.pdf (accessed: June, 2012)
18. Java OpenGL Game Engine: jMonkeyEngine (April 2012), http://jmonkeyengine.com/ (accessed: April, 2012)
19. FreeIMU library, http://www.varesano.net/topic/freeimu (accessed: April, 2012)

Study of a Wake Up Radio Architecture for Home Multimedia Networks

Aissa Khoumeri, Florin Hutu, Guillaume Villemaud, and Jean-Marie Gorce

Abstract. A theoretical study on the impact of using a wake-up radio architecture in terms of energy consumption is proposed. The main objective is to reduce the overall energy consumption of a home multimedia networks. The energy consumed by the proposed wake-up radio architecture is compared to a classical WiFi architecture, for an ad-hoc scenario. The sleep time has an important role to compare the dissipated energy. This study demonstrates that the longer the sleep time the better the energy saved is obtained by the wake-up architecture.

1 Introduction

Nowadays wireless local area networks (WLANs) are widely deployed to provide internet access. Even at home, the number of devices to be connected to the internet increases and consequently the energy consumption of the home multimedia network grows. The necessity to reduce the energy consumption of the connected devices is an ongoing challenge for researchers in order to develop a green home wireless network. Worldwide consortiums gathering academic and industrial partners are formed and gave themselves as challenges to increase the energy efficiency by a factor of 1000 in the next few years [1].

Aissa Khoumeri · Florin Hutu · Guillaume Villemaud · Jean-Marie Gorce
Université de Lyon, INRIA, INSA-Lyon, CITI-INRIA,
6, avenue des Arts, F-69621, Villeurbanne
e-mail: {aissa.khoumeri,floin-doru.hutu,Guillaume.Villemaud,
jean-marie.Gorce}@insa-lyon.fr

L. De Strycker (ed.), *ECUMICT 2014,*
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_6, © Springer International Publishing Switzerland 2014

An approach to reduce the energy consumption of the home multimedia network is to shut down the radio-frequency (RF) part of the data interfaces when there is no communication demand, for example the "power save mode" used in IEEE 802.11 standard [2].

The energy consumption of the radio interface in the idle mode remains relatively high because of the low energy efficiency of the RF front-ends. For example, in the case of Linksys WRG54G Wi-Fi access point (AP), the maximum reduction in the energy consumption that can be achieved is 28.57 % [3]. This value represents the energy consumption of Wi-Fi interface compared to the entire energy consumption of the access point.

The wake-up concept has been introduced in the context of wireless sensor networks (WSN) as a way to optimize the battery lifetime [4] [5] [6] [7]. In WSN, the duty-cycle MAC protocols to reduce the overall power consumption exist. Wireless sensors in a duty cycle mechanism, should periodically sleep and wake their radio modules (like S-MAC, T-MAC, and B-MAC [8] [9] [10] [11]). High duty cycle reduces the latency but increases the energy consumption, while low duty cycle reduces the energy consumption but increases the latency. An appropriate trade-off between energy consumption and latency has to be achieved.

Another wake-up mechanism radio is to use a secondary low power transceiver to monitor the channel instead of the main data transceiver. In this paper, the wake-up architecture designs both the secondary low power transceiver and the main one.

Usually, the On Off Keying (OOK) modulation and the non-coherent envelope detection technique are used to design the low power secondary transceivers. This is because of their simplicity to implement, and their low energy consumption.

The use of a wake-up radio in the home wireless network, like personnel digital assistant (PDAs), has been firstly presented in [12]. In perspectives of this work [3], a low power sleep mode and an out of band wake-up mechanism has been adopted. A low power radio module to carry out of band control information is used to switch the WiFi access point (AP) into the sleep mode when no users exist.

The use of secondary low power wireless module, which has the same frequency and shares the same antenna with the co-located WiFi radio interface, was proposed in [13]. Based on modulation of the IEEE 802.11g frame length, the wake-up signal recognition needs an analogue to digital converter and a signal detection module that introduces more power consumption to the wake-up receiver.

The authors previous works [14] [15], present a wake-up radio architecture based on a frequency pattern identification of the address equipments. The secondary low power wireless module does not need baseband treatment of the identifier, and the decision to wake up is taken directly in the analog part. This implies a reduction of the energy consumption in sleep mode and a reduction of latency. In this paper an evaluation of benefits of using such architecture in terms of energy consumption is given.

The rest of this paper is organized as follows. Section 2 presents the proposed wake-up architecture and the identification technique. Section 3 presents the energy consumption estimation in an ad-hoc scenario. Finally, section 4 presents some conclusions and perspectives of this work.

2 Wake-Up Radio Architecture Description

The proposed wake-up architecture in [14] [15] has an ultra-low power secondary wake-up radio circuit that monitors the channel for the wake-up signal identifier. In order to minimize the energy consumption, the wake-up receiver is designed to be simple as possible, as seen in figure 1.

The wake-up circuit has two paths in which the input signal passes through, (i) the direct path which contains the multi-band filter that has the same bandwidth as the spectrum formed by the identifier signal (as the one presented in figure 2) and (ii) the complementary path contains the complementary multi-band filter, which has its pass bandwidth situated in the band stop of the direct filter. This configuration allows the wake-up radio circuit to eliminate the interference signal generated by other equipment present in the same area.

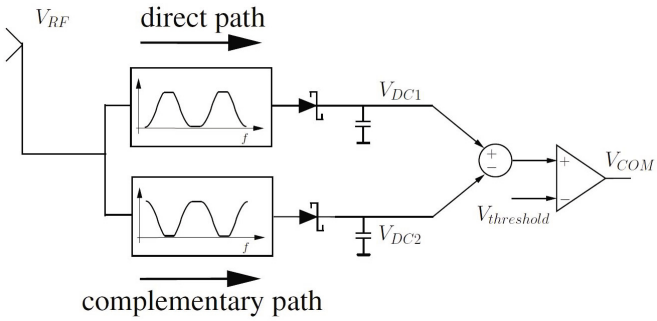


Fig. 1 Wake-up radio circuit

The identifier is formed with an arrangement of a WLAN power spectrum, obtained by a selection of specific OFDM sub-carriers of the same group G_i in the channel. figure2 shows an example of an identifier power spectrum. The 64 sub-carriers in the total band of 20 MHz are divided into 4 groups. Each group contains 14 neighbor sub-carrier of 312.5 KHz as shown below:

$$\begin{cases} G_1 = \{-28...-15\} \\ G_2 = \{-14...-1\} \\ G_3 = \{1...14\} \\ G_4 = \{15...28\} \end{cases} \quad (1)$$

Because of the complementary configuration of the architecture, the same DC voltage is obtained in the two paths when another signal is received for example the WLAN RF signal, in this case the difference of the two DC voltages is null.

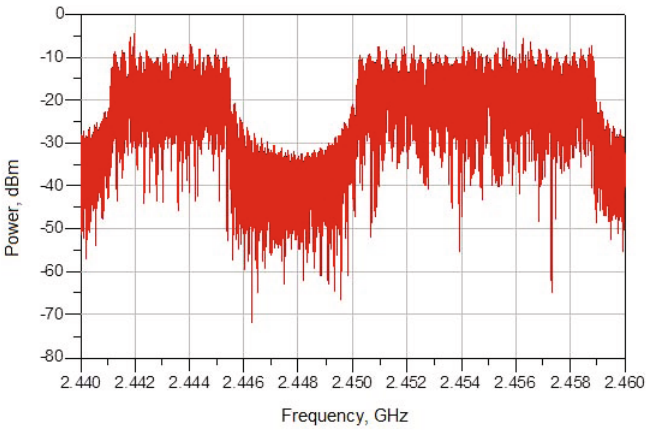


Fig. 2 Power spectrum of an identifier G_1 =ON G_3 =ON and G_4 =ON

When the identifier is received, the output DC voltage VDC1 obtained from the rectifier has the maximum value and the DC voltage VDC2 is null. In this case, the difference between the two DC voltages is higher than $V_{threshold}$, and the activation signal V_{COM} is generated as shown in figure 3.

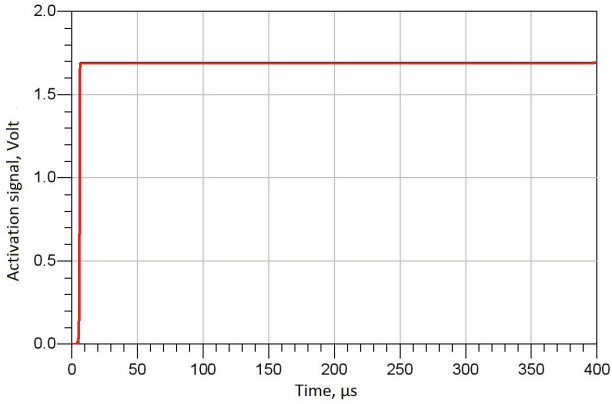


Fig. 3 Wake-up signal activation V_{COM}

3 Energy Consumption Estimation

In order to calculate the energy consumption in a wake-up radio architecture, we need to consider the energy consumption of the wake up circuit and the main data radio interface in different communication states. In this study an ad-hoc scenario as shown in figure 4 is considered. Equipment 2 has its main interface "Main data 2" off, and its secondary wake-up radio receiver monitors the channel. When the wake-up signal identifier is sent by the equipment 1, then the equipment 2 passes from offline state to the receiving state, and the data transfer between the two main interfaces may begin.

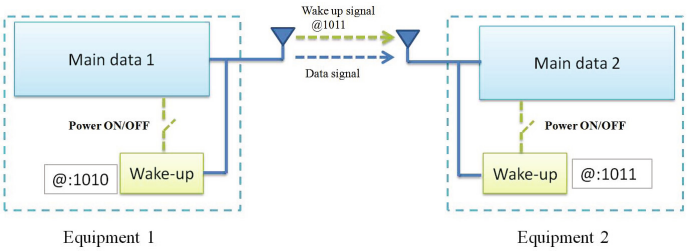


Fig. 4 Proposed Wake-up radio architecture in an ad-hoc scenario

The main data interface used in the study is the 2.4-2.5 GHz transceiver MAX2830 [16]. Figure 5 gathers the different timing state and power consumption of this front-end. This timing state starts when the wake-up radio circuit receives during 400 μs the identifier signal.

In the case of a classical WiFi the MAC layer controls the state of the physical layer (PHY). This is achieves by following the distributed coordinated function (DCF) access method [2]. All stations (STA) that intend to transmit frames must monitor the channel to determine if other STA are transmitting.

If the channel is idle for an interval of time that exceeds the distributed interframe space (DIFS), then the packet is transmitted. Otherwise, the STA monitor the channel until it is sensed idle for a DIFS interval. After that it generates a random backoff interval chosen in the range of [0,CW], where CW is the minimum contention window (CW_{min}). Next the STA will reduce this number by one each time an idle slot is elapses, or it is frozen when the medium is sensed busy.

The energy consumption E_{WLAN}^{active} to transmit and receive one frame by the classical WiFi architecture is given by Equation (2). Also the energy consumed in the case of a wake up radio architecture $E_{Wake-up}^{active}$ is given by Equation (3).

Table 1 defines the relevant parameters used in the energy model. These parameters are taken from the MAC layer of a standard 802.11g, based on OFDM PHY, and the power consumption in different states (transmitting, receiving, idle and sleep) of the MAX2830 transceiver.

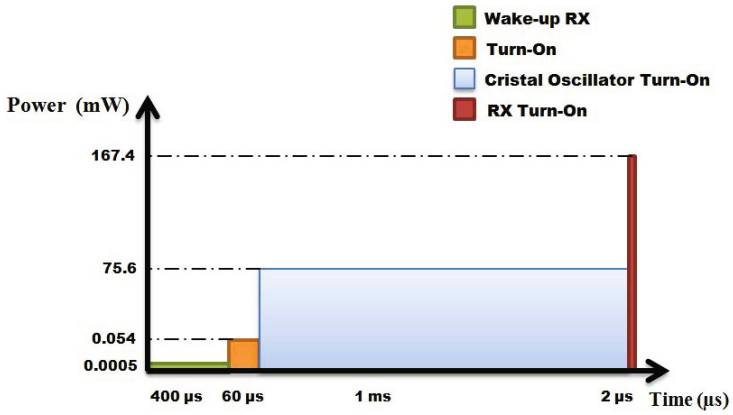


Fig. 5 Time and energy spend to wake up with MAX2830

Equation (4) defines the energy consumed by the classical WiFi for sending and receiving one data frame for a classical WiFi architecture :

$$E_{WLAN}^{active} = (P_{TX} + P_{RX}) \times (T_{data} + T_{ACK}) + P_{idle} \left(2 \times T_{SIFS} + T_{DIFS} + T_{Backoff}\right) \quad (2)$$

Equation (5) defines the energy consumed in the wake-up architecture for sending and receiving one data frame :

$$E_{Wake-up}^{active} = (E_{WLAN}^{active} - P_{idle} \left(T_{DIFS} + T_{Backoff}\right)) + \left(P_{TX} + P_{wake-up}\right) T_{wake-up} + P_{Turn-On} \times T_{Turn-0n} + P_{oscil} \times T_{oscil} + P_{RX} \times T_{RX-On} \quad (3)$$

The energy consumed in sleep mode by the classical WiFi architecture is given by :

$$E_{WLAN}^{sleep} = 2 \times P_{sleep} \times T_{sleep} \quad (4)$$

The energy consumed in sleep mode by the wake-up architecture is given :

$$E_{Wake-up}^{sleep} = 2 \times P_{wake-up} \times T_{sleep} \quad (5)$$

The total energy consumed by the classical WiFi is given by :

$$E_{WLAN}^{total} = E_{WLAN}^{active} + E_{WLAN}^{sleep} \quad (6)$$

The total energy consumed by the wake-up architecture is given by :

$$E_{Wake-up}^{total} = E_{Wake-up}^{active} + E_{Wake-up}^{sleep} \quad (7)$$

Table 1 Parameters for modeling the energy consumption using a MAX2830 802.11g transceiver

Variable	description	value
P_{TX}	Power consumed in transmitting mode	800 mW
P_{RX}	Power consumed in reception mode	167.4 mW
P_{idle}	Power consumption in standby mode	75.6 mW
P_{sleep}	Power consumed in sleep mode	75.6 mW
$P_{wake-up}$	Power consumed by the wake up secondary circuit	500 nW
$P_{Turn-On}$	Power consumed by the radio interface in shutdown mode	54 μW
P_{oscil}	Power consumed by the frequency synthesizer block	75.6 mW
T_{data}	Time spend for sending 1500 byte of data at 9 Mbps	2.42 ms
T_{SIFS}	Time of short interframe space	10 μs
T_{DIFS}	Time of distributed interframe space	28 μs
$T_{Backoff}$	Time of backoff	67.5 μs
$T_{wake-up}$	Time of wake up identifier	400 μs
$T_{Turn-0n}$	Time to pass from off mode to shutdown mode	60 μs
T_{RX-On}	Time to pass from standby mode to receive mode	2 μs
T_{sleep}	Time spend in sleep state	Variable

Figure 6 shows the energy consumption of the classical WiFi, Equation (6), as well as the energy consumption of the wake-up architecture, Equation (7), when sending and receiving one frame of 1500 bytes with a rate of 9 Mb/s.

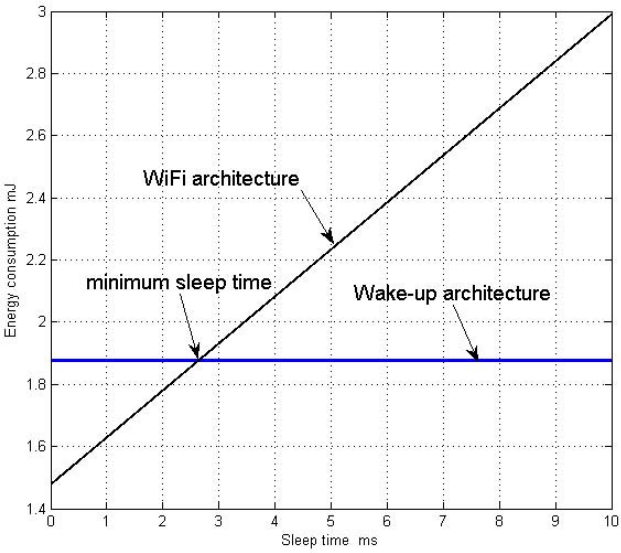


Fig. 6 Energy consumption vs sleep time

In Equation (8), T_{sleep}^{min} represents the sleep time where the wake-up architecture has the same energy consumption as a classical WiFi architecture. From Equations (6) and (7) we get the minimum sleep time T_{sleep}^{min}:

$$T_{sleep}^{min} = \frac{1}{2} \times \frac{E_{Wake-up}^{total} - E_{WLAN}^{total}}{P_{sleep} - P_{wake-up}} \tag{8}$$

If the sleep time is less than $T_{sleep}^{min} = 2.6\ ms$, the energy consumption of the wake-up architecture is higher than the energy consumption of the classical WiFi architecture. This is due to the energy consumed by the main data when switching from offline to the receiving mode.

When the sleep time is higher than T_{sleep}^{min}, the energy consumption of wake-up architecture is less than a classical WiFi architecture, because of it low energy consumption in sleep mode $(500\ nW)$.

In figure 7 the energy consumption in function of the sleep time at different data rate is plotted. The figure shows that when the enhancement condition is respected $\left(T_{sleep}^{min} > T_{sleep}\right)$, the higher is the sleep time, the better is the enhancement.

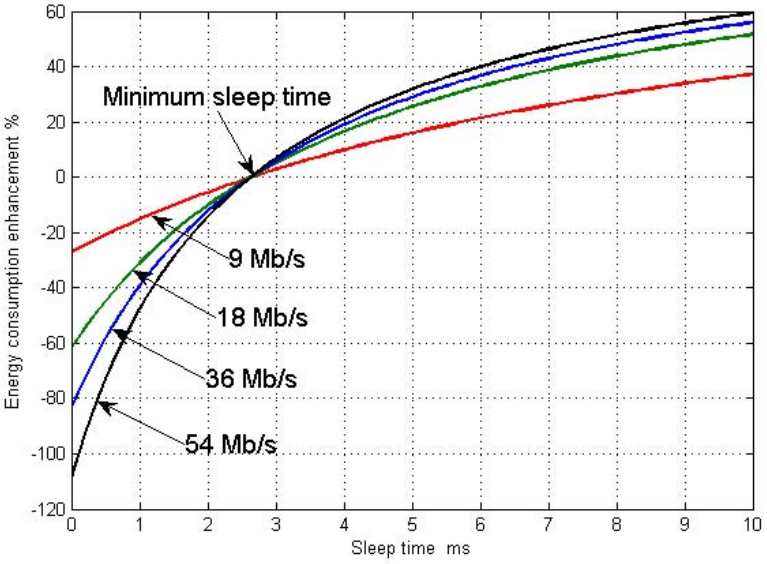


Fig. 7 Energy consumption enhancement vs sleep time

Indeed when the sleep time is equal to 10 *ms*, the wake-up architecture achieves up to 40 % (*resp* 60 %) of energy saving compared to a classical WiFi architecture at 9 Mb/s (*resp* 54 Mb/s) data rate. Moreover, we notice that the higher is the

data rate better is the energy saving, this mainly due to the transmission time of the packet which is inversely proportional to the data rate.

4 Conclusion

In this paper, the energy consumption of a proposed wake-up radio architecture has been evaluated and compared to a classical WiFi. This study demonstrates that, with respect to minimum sleep time T_{sleep}^{min}, the wake-up architecture saves more energy than the classical WiFi. Future works include analyzing a network scenario, with a traffic rate (on the order of hours) to find the maximum sleep time in different traffic loads. However in such scenario, latency and false wake up should be studied in order to quantify the robustness of the proposed architecture. Moreover, a prototype will be manufactured in order to have concludent measurements of energy consumption and time latency.

References

1. GreenTouch Consortum online, `http://www.greentouch.org/`
2. IEEE standard for information technology and information exchange between systems - local and metropolitan area networks - specific requirements - part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications. IEEE Std 802.11, pp. 1–1238 (2007)
3. Haratcherev, I., Fiorito, M., Balageas, C.: Low-power sleep mode and out-of-band wake-up for indoor access points. In: IEEE GLOBECOM Workshops, 2009, pp. 1–6. IEEE (2009)
4. Gu, L., Stankovic, J.A.: Radio-triggered wake-up for wireless sensor networks. Real-Time Systems 29(2-3), 157–182 (2005)
5. Ansari, J., Pankin, D., Mähönen, P.: Radio-triggered wake-ups with addressing capabilities for extremely low power sensor network applications. International Journal of Wireless Information Networks 16(3), 118–130 (2009)
6. Durante, M.S., Mahlknecht, S.: An ultra low power wakeup receiver for wireless sensor nodes. In: Third International Conference on Sensor Technologies and Applications, SENSORCOMM 2009, pp. 167–170. IEEE (2009)
7. Khoumeri, A., Hutu, F., Villemaud, G., Gorce, J.-M.: Wake-up radio architectures used in wireless sensor networks. In: COST IC1004, Lyon, France (May 2012)
8. Ye, W., Heidemann, J., Estrin, D.: An energy-efficient mac protocol for wireless sensor networks. In: Proceedings of the IEEE Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies, pp. 1567–1576. IEEE (2002)
9. Van Dam, T., Langendoen, K.: An adaptive energy-efficient mac protocol for wireless sensor networks. In: Proceedings of the 1st International Conference on Embedded Networked Sensor Systems, pp. 171–180. ACM (2003)
10. Polastre, J., Hill, J., Culler, D.: Versatile low power media access for wireless sensor networks. In: Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems, pp. 95–107. ACM (2004)
11. Ullah, N., Khan, P., Kwak, K.S.: A very low power mac (vlpm) protocol for wireless body area networks. Sensors 11(4), 3717–3737 (2011)

12. Shih, E., Bahl, P., Sinclair, M.J.: Wake on wireless: an event driven energy saving strategy for battery operated devices. In: Proceedings of the 8th Annual International Conference on Mobile Computing and Networking, pp. 160–171. ACM (2002)
13. Tang, S., Yomo, H., Kondo, Y., Obana, S.: Wake-up receiver for radio-on-demand wireless lans. EURASIP Journal on Wireless Communications and Networking 2012(1), 1–13 (2012)
14. Khoumeri, A., Hutu, F.D., Villemaud, G., Gorce, J.-M.: Proposition d'une architecture de réveil radio utilisée dans le contexte des réseaux multimédia domestiques. In: Journées Nationales Microondes, Paris, France (May 2013)
15. Hutu, F., Khoumeri, A., Villemaud, G., Gorce, J.-M.: Wake-up radio architecture for home wireless networks. In: 2014 IEEE Radio and Wireless Symposium (RWS), Newport Beach, USA (in press, January 2014)
16. MAXIM, MAX2830 Industry's 2.4GHz to 2.5GHz 802.11g/b RF Transceiver, PA, and Rx/Tx/Antenna Diversity Switch (April 2008)

Hybrid Multi-objective Network Planning Optimization Algorithm

Ning Liu, David Plets, Wout Joseph, and Luc Martens

Abstract. In this paper, a novel hybrid algorithm for the optimization of indoor wireless network planning is applied to a polyvalent arts centre. The results of the algorithm are compared with those of a heuristic network planner for three scenarios. Results show that our algorithm is effective for optimization of wireless networks, satisfying maximum coverage, minimal power consumption, minimal cost, and minimal human exposure.

1 Introduction

When planning wireless networks, different characteristics of the result can be considered and optimized, e.g. coverage, energy consumption, exposure and cost. In [5], energy conservation techniques on different types of base stations were compared. Exposure in office environments has been investigated in [8] and [21]. As for wireless network planning optimization with four main requirements, in [18, 20], researchers have focused on femtocells and hybrid (DVB-H/UMTS) networks, since these networks are associated with improved coverage and lower exposure. Plets et al. have presented a heuristic to optimize the exposure in indoor wireless networks, which is named the WiCa Heuristic Indoor Propagation Prediction (WHIPP) tool [16, 14, 15].

Ning Liu
School of Computer Science and Engineering,
University Electronic Science and Technology of China, Chengdu, China
e-mail: ning.liu@intec.ugent.be

Ning Liu · David Plets · Wout Joseph · Luc Martens
WiCa, Ghent University / iMinds, Dept. of Information Technology,
Gaston Crommenlaan 8 box 201, B-9050 Ghent, Belgium
e-mail: {ning.liu,david.plets,wout.joseph,
luc.martens}@intec.ugent.be

L. De Strycker (ed.), *ECUMICT 2014,*
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_7, © Springer International Publishing Switzerland 2014

Mainly three types of optimization algorithms are considered when optimizing indoor wireless environments planning [22, 13, 23]: PSO (Particle Swarm Optimization) [8], ACO (Ant Colony Optimization) [21] and GA (Genetic Algorithm) [18]. In [11], researchers use ACO to optimize the wireless networks in order to achieve coverage in energy-efficient way. In [3], Chen proposed an altered version of the PSO algorithm to solve the network planning problem in RFID systems. GAs have been developed to plan wireless communication networks in [9, 24] and have also shown good performance for coverage optimization and exposure minimization in [10, 2]. GA and PSO algorithms have both yielded successful results and fast convergence in this field [22, 23], while ACO needs much more iterations for optimizing wireless network in [23].

In [12], a hybrid algorithm (combining GA and quasi-PSO) was proposed for the optimization of the wireless network planning, accounting for four requirements: maximum coverage, minimal power consumption, minimal cost, and minimal human exposure. In this paper, this algorithm and the WHIPP algorithm will be applied to a polyvalent arts centre for three different optimization scenarios. Section 2 briefly introduces the configuration and the fitness functions that are used. In Section 3, three scenarios are presented. A summary of our hybrid algorithm is provided in Section 4. The results and comparison with WHIPP of these scenarios for the indoor environment are provided in Section 5. Conclusions are presented in Section 6.

2 Configuration and Fitness Function

2.1 Configuration

Fig. 1 shows a map of the ground floor of the Vooruit cultural centre (a polyvalent arts center). It is mainly constructed with large concrete walls and glass. The goal is to design a wireless network with WiFi (801.11n) access points operating at a frequency of 2.4 GHz, with an antenna gain of 2 dBi, and for required received power of -68 dBm (for HD video coverage). The EIRP (Effective Isotropic Radiated Power) range of the access points runs from 0 to 20 dBm. The receiver antenna gain is 0 dB. There are 202 possible positions to place WiFi access points; these are also the receiver points for which coverage and exposure will be calculated. The path loss PL (the ratio of the transmitted power and the received power) will be modeled according to the following two models.

- The first model is the two-slope model proposed by the IEEE 802.11 **TGn channel models** group [7].

$$\begin{aligned} PL(d) &= PL_{free}(d) + X \ \ (d \leq d_{br}) \\ PL(d) &= PL_{free}(d) + 32log_{10}\left(\frac{d}{d_{br}}\right) + X \ \ (d > d_{br}) \end{aligned} \tag{1}$$

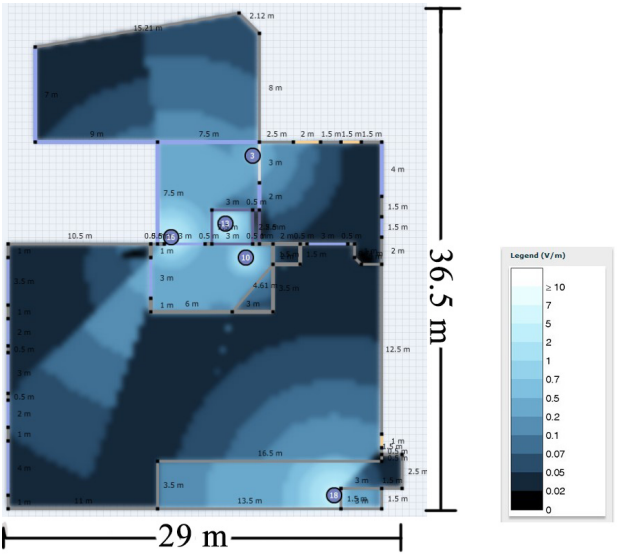


Fig. 1 Map of The Indoor Environment and the Exposure Level for Scenario III for the SIDP Model

Where $PL_{free}(d)$ is the free space loss [19]. The variation of path loss X due to shadowing follows a lognormal distribution, with two different standard deviations σ [dB] of X for $d \leq d_{br}$ and $d > d_{br}$. In this situation, parameters are considered as follows: d_{br} of $10m$, $\sigma = 3\ dB$ for $d \leq d_{br}$ and $\sigma = 5\ dB$ for $d > d_{br}$, corresponding to a 95% shadowing margin of $4.92\ dB$ and $8.2\ dB$ for $d \leq d_{br}$ and $d > d_{br}$ respectively [7]. The temporal fading margin is set at $5\ dB$ [1].

- The second model is **Simple Indoor Dominant Path Loss model** used in [17]. The shadowing margin is set at $7\ dB$ (95%) and the fading margin at $5\ dB$ (99%).

2.2 Fitness Functions

Four different fitness functions will be investigated for the optimization of the network planning. Each fitness function optimizes one or more of the four main wireless network characteristics (coverage, power consumption, cost, human exposure). The results of the different functions, f_i $(i = 1, 2, 3)$ will range from 0 to 100, so that they have an equal contribution when they are combined in a new fitness function (see Section 2.2.4). A comparable value of the weights (w_1, w_2, w_3) of the different functions (f_1, f_2, f_3) then causes a comparable influence of the function on the combined fitness function (f_4).

2.2.1 Coverage

The first fitness function represents coverage fitness as in Eq. (2),

$$f_1 = 100\frac{f_{sol}}{f_{tot}} \tag{2}$$

Where f_{tot} is the number of all reception points (202 for the considered building), f_{sol} is the number of reception points covered by the current solution in this indoor environment and f_1 represents the coverage percentage of the considered network configuration.

2.2.2 Power Consumption and Economic Cost

In Eq. (3), f_2 represents the ratio of the actual power consumption of the considered network configuration to the maximum achievable power consumption in the network:

$$f_2 = 100\frac{\sum_{i=1}^{n} p_i}{p_{max}}, \tag{3}$$

where p_i is the power consumption of the i-th access point (12 W for a WiFi access point which is on [6], 0 W when it is turned off), p_{max} is the total power consumption when all possible access points are turned on. The actual EIRP also affects the total power consumption. However, because the impact is small [6], we neglect the effect of the radiated power and assume a fixed value of 12W per access point [4, 6]. Eq. (3) then reduces to

$$f_2 = 100\frac{m}{n}, \tag{4}$$

where m is the number of access points which are turned on, and n total number of possible positions (202 for Vooruit).

The cost of all installed access points represents the economic cost (Capital Expenditures). Since a fixed power consumption is assumed for all access points, f_2 represents both the economic cost fitness function and the power consumption fitness function of the considered network deployment.

2.2.3 Exposure

In Eq. (5), f_3 is a fitness function based on the median electric-field strength $E_m\,[V/m]$ observed at the considered receiver points in the environment.

$$f_3 = 100\frac{E_m}{E_{max}}, \tag{5}$$

where E_{max} is the maximal median electric-field value that could be achieved. This is the case when all (202) access points are turned on with an EIRP of 20 dBm,

yielding a value for E_{max} of 2.19 V/m is obtained for TGn model and 2.46 V/m for SIDP model. The optimal solution of this fitness function has a minimal median electric field strength.

2.2.4 Combined Fitness Function

In Eq. (6), f_4 is a global fitness function which combines above three presented fitness function:

$$f_4 = w_1 f_1 - w_2 f_2 - w_3 f_3 \tag{6}$$

where w_i is the weight (values between 0 and 1) of function f_i with its value determined by the importance of f_i. By adjusting w_i , four demands can be jointly optimized. 2. w_2 and w_3 are chosen so that coverage is the most important factor in optimization ($w_1 = 1$). However, on top of coverage optimization, energy consumption (w_2) and exposure (w_3) are also important, but less than coverage. The values of w2 and w3 need to be small enough to obtain a solution with 100% coverage, but large enough to still minimize energy consumption and exposure. Consequently, when we increase w2, results with less access points are expected. When we increase w_3, results with lower exposure levels are expected. The weights control the value of the fitness function and the fitness value affects the result of the algorithm. The best solutions are the ones with the highest combined fitness function values, as they correspond to higher coverage rates, lower total power consumptions (and cost), and lower exposure values. For the optimization of the fitness function, a hybrid genetic optimization algorithm is used [12].

3 Scenarios

We define three scenarios to investigate the influence of coverage, exposure, and cost restrictions on the network deployment for Vooruit (in Fig. 1) by applying our algorithm and comparing with the WHIPP algorithm. Unlike for our hybrid optimization algorithm, the WHIPP optimization is not based on the use of a fitness function and the evaluation of a number of iterations. It builds a solution based on a number of optimization phases following a fixed procedure. The WHIPP algorithm allows an optimization for 100% coverage with a minimal number of APs, as well as an optimization for 100% coverage with a minimal exposure. This allows a comparison with the output of Scenarios I and II by our algorithm, as described hereafter. All scenarios are applied to the configuration and using the PL model of Section 2.

3.1 Scenario I: Coverage with Minimal Number of APs

Scenario I aims to obtain 100% coverage rate with a minimal number of access points (minimal both cost and power consumption). We select the weight w_2 for the f_2 as 0.2, since this value is large enough to minimize the number of APs and small enough to aim for a coverage rate of 100%. The combined fitness function of Eq. (6) in scenario I is as follows:

$$f_4 = f_1 - 0.2f_2 \quad (w_1 = 1, w_2 = 0.2) \tag{7}$$

3.2 Scenario II: Coverage with Minimal Human Exposure

Scenario II intends to obtain 100% coverage rate with a minimal median exposure. The combined function f_4 is as follows:

$$f_4 = f_1 - 0.2f_3 \quad (w_1 = 1, w_3 = 0.2) \tag{8}$$

We select the weight for the exposure level fitness w_3 as 0.2, since this value is large enough to minimize the exposure level and small enough to obtain 100% coverage.

3.3 Scenario III: Coverage with Minimal Human Exposure and Minimal Number of APs

Scenario III is defined to consider a tradeoff among a high coverage rate, a low total power consumption and a low median electric-field strength. Under the condition of scenario III in Eq. (9), we consider different requirements together: coverage, number of access points (cost and power consumption), and exposure level.

$$f_4 = f_1 - 0.2f_2 - 0.2f_3 \quad (w_1 = 1, w_2 = 0.2, w_3 = 0.2) \tag{9}$$

4 Our Algorithm

Fig. 2 shows the flow chart that corresponds to the operation of our algorithm. The main operations of the genetic algorithm are crossover and mutation operations.

Firstly, 1000 random solutions are generated and their fitness values are calculated. The solutions with the top-80 fitness values are put into a list.

Secondly, after sorting this solution list based on their fitness values, the top-40 of the list with the high fitness values is called 'good list' and the rest of the list is called 'bad list'.

Thirdly, new solutions are generated by a crossover operation between a father solution from the good list and a mother solution from the bad list. In this operation one third of the father solution is combined with two third of a mother solution. If the offspring gets a higher fitness value than that of mother solution, we put it into the corresponding location of the list.

Fourthly, the mutation operation adds random changes in a solution and makes the algorithm converge to a global optimum instead of to a local optimum. During each mutation, a solution has equal possibility to perform one of the following operations:

- Turn off one access point in this solution;
- Turn on an access point with random power value;
- Turn on an access point with random power value and turn off another access point of this solution;
- Turn off two access points of this solution and turn on an access point with random power value;
- Change the power value of an access point of this solution;
- Change the position of an access point of this solution.

GAs and PSOs are suitable to solve the multi-objective problem described in Section 1. Since we can obtain benefit from the evaluation and heredity of GA, the GA is better than PSO. PSO performs better, when the solution consists of only one AP, due to a slight change of solution in each iteration is better to quickly find the optimal solution. Therefore, our algorithm introduces operations of PSOs into the GA system. When only one access point is sufficient, offspring are generated by using the quasi-PSO with a certain probability. The new algorithm approaches the global optimum more efficiently.

5 Results

5.1 Simple Indoor Dominant Path Loss Model (SIDP)

The results for the scenarios described above are investigated for WHIPP and our algorithm based on the SIDP model and are listed in Table 1. For all scenarios, the coverage of all methods is 100%. For scenario I, our algorithm obtains a solution with 3 access points, while WHIPP obtains a solution with 4 access points. The solution of our algorithm generates a lower median exposure level of 123.7 mV/m versus 155.6 mV/m of WHIPP, due to the lower number and EIRP of APs of the solution of our algorithm. The solution of our algorithm for scenario II also generates a slightly lower median exposure level (9.3% lower) than that of WHIPP, although, the 95% percentile exposure level of our algorithm is much higher than that of WHIPP, since there is less spatial homogeneity in the exposure levels of our solution. The solution of our algorithm for scenario III is a compromise between all criteria (high coverage, low exposure and needs a low number of APs). It shows

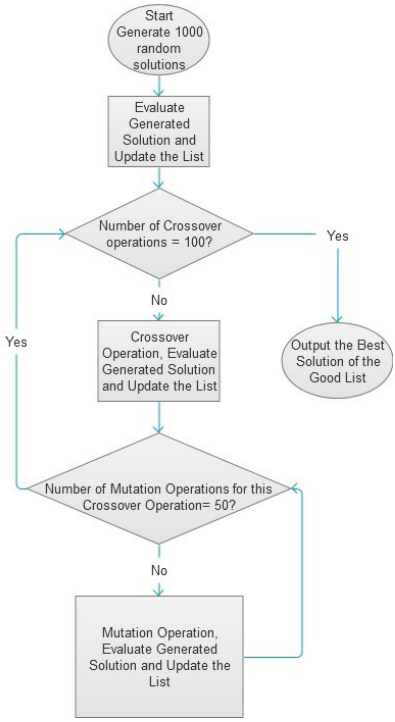


Fig. 2 Flow Chart of Our Algorithm

the advantage of our algorithm, since scenario III is difficult to implement in the WHIPP tool. The solution of our algorithm for scenario III requires 5 APs and generates a median exposure level of only 47.2 mV/m which is about 1% higher than that of scenario II which needs 10 APs. Fig. 1 shows the electronic field distribution for scenario III in the considered building. The location and EIRP of the APs is also indicated. Compared to WHIPP [15, 14], the simulation time (last column in Table 1) of the our algorithm is always much higher than that of WHIPP, since WHIPP is a heuristic. Limiting the simulation time of our algorithm to the WHIPP simulation times would yield worse results, since a substantial number of iterations is required for this type of algorithms (GAs). However, since network planning is mostly a task that is performed only once, large computation times are not really an issue if the algorithm finally provides a better result.

Fig. 3 shows the comparison of CDF of the exposure values based on the SIDP model for WHIPP and our algorithm. It shows that the exposure level of our algorithm is always lower than that of WHIPP at the same probability for scenario I. However, when we consider scenario II, this situation is reversed when the probability greater than 80% (see Fig. 3), since the less spatial homogeneity in the exposure

Table 1 The Results of Scenarios for indoor Environment Based on Simple Indoor Dominant Path Loss Model

Scenarios	Method	Coverage Rate [%]	#APs [-]	E_{50}[a] $[mV/m]$	E_{95}[b] $[mV/m]$	EIRP $[dBm]$	Simulation Time [s]
Scenario I	WHIPP	100	4	155.6	819.7	$4 \times 20dBm$	111
	Our Algorithm	100	3	123.7	775.0	$15dBm, 2 \times 20dBm$	8.8×10^3
Scenario II	WHIPP	100	12	51.6	190.2	$-26dBm$, $-13dBm$, $-1dBm$, $0dBm$, $5 \times 1dBm$, $2dBm$, $4dBm$, $5dBm$	274
	Our Algorithm	100	10	46.8	422.2	$2 \times 0dBm$, $2dBm$, $4dBm$, $2 \times 5dBm$, $3 \times 9dBm$, $17dBm$	7.2×10^4
Scenario III	WHIPP[c]	-	-	-	-	-	-
	Our Algorithm	100	5	47.2	465.3	$3dBm$, $10dBm$, $13dBm$, $16dBm$, $18dBm$	6.6×10^4

[a] E_{50}: 50% percentile of E (mV/m)
[b] E_{95}: 95% percentile of E (mV/m)
[c] WHIPP cannot optimize 3 requirements as required for scenario III

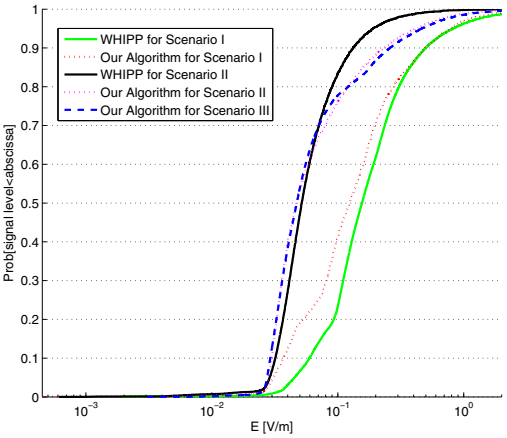


Fig. 3 Comparison The CDF of The Exposure Results for Indoor Environment (Based on The Simple Indoor Dominant Path Loss Model)

levels of our solution. The exposure level of our algorithm for scenario III is very close to that of our algorithm for scenario II.

Table 2 The Results of Scenarios for Indoor Environment Based on TGn Two-Slope Path Loss Model

Scenarios	Method	Coverage Rate [%]	#APs [-]	E_{50}[a] [mV/m]	E_{95}[b] [mV/m]	EIRP [dBm]	Simulation Time [s]
Scenario I	WHIPP	100	2	164.1	631.8	$2 \times 20dBm$	1
	Our Algorithm	100	2	115.5	434.9	$8dBm$, $19dBm$	35
Scenario II	WHIPP	100	6	35.0	116.1	$6 \times 1dBm$	6
	Our Algorithm	100	5	34.5	118.6	$2 \times 1dBm$, $3 \times 2dBm$	137
Scenario III	WHIPP[c]	-	-	-	-	-	-
	Our Algorithm	100	4	41.6	294.0	$1dBm$, $2 \times 2dBm$, $18dBm$	104

[a] E_{50}: 50% percentile of E (mV/m)
[b] E_{95}: 95% percentile of E (mV/m)
[c] WHIPP cannot optimize 3 requirements as required for scenario III

5.2 TGn Model

Table 2 lists the results of WHIPP and our algorithm. As for scenario I, WHIPP and our algorithm both obtain a solution with 2 APs. The median and the 95% percentile exposure levels of our algorithm for scenario I are both lower than that of WHIPP, due to the lower EIRP of the APs of our algorithm. The differences between the exposure results of WHIPP and our algorithm for scenario II is small. The solution of our algorithm needs 6 APs, while that of WHIPP needs 5 APs. For the exposure level for scenario II, the median exposure level of WHIPP is 1.5% higher than that of our algorithm. However, E_{95} of our algorithm is 2.1% higher than that of WHIPP. For scenario III (Table 2), our algorithm obtains a solution with 4 APs (20% lower than that of our algorithm for scenario II) and generates a median exposure of 41.6 mV/m (74.6% lower than that of our algorithm for scenario I). As for the simulation time, that of WHIPP is again always lower than that of our algorithm for each scenario, but calculation times are limited for a algorithm as well (maximum =137s for scenario III).

Comparison of the CDFs for the TGn model shows that the exposure values for our algorithm are mostly lower than for WHIPP at the same probability for scenario I (see Fig.4). The difference between the exposure levels of WHIPP and that of our algorithm for scenario II is small. For scenario III, the curve of our algorithm is between scenario I (minimal cost or number of APs) and scenario II (minimal exposure).

6 Conclusions

A hybrid genetic optimization algorithm has been proposed to optimize coverage rate, human exposure to radio-frequency sources, energy consumption and economic

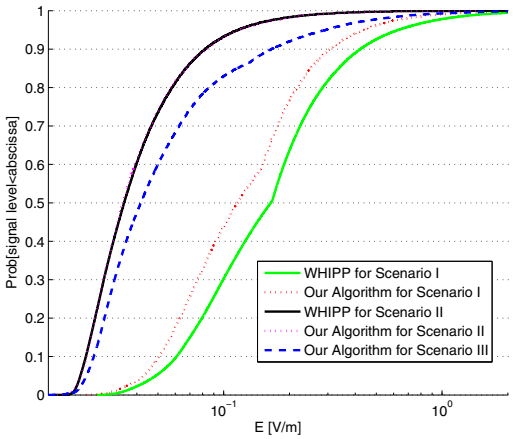


Fig. 4 Comparison The CDF of The Exposure Results for Indoor Environment (Based on The TGn Model)

cost of the indoor wireless networks. Specific fitness functions were used to evaluate the solutions for a homogeneous WiFi network. Three scenarios are defined to verify the performance of the algorithm and good results are obtained. An application for a realistic indoor environment (Vooruit) is investigated leading to reductions of cost and exposure when applying our algorithm compared to a heuristic tool (a median exposure level reduction of 9% or a cost reduction of 25% are obtained compared to WHIPP based on the SIDP model). Future research enable planning of heterogeneous wireless networks for various indoor environments.

References

1. Bultitude, R.J.C.: Measurement, characterization and modeling of indoor 800/900 mhz radio channels for digital communications. IEEE Communications Magazine 25(6), 5–12 (1987), doi:10.1109/MCOM.1987.1093629
2. Cerri, G., De Leo, R., Micheli, D., Russo, P.: Base-station network planning including environmental impact control. IEE Proceedings-Communications 151(3), 197–203 (2004), doi:10.1049/ip-com:20040146(410)151
3. Chen, H., Zhu, Y., Hu, K., Ku, T.: RFID network planning using a multi-swarm optimizer. Journal of Network and Computer Applications 34(3), 888–901 (2011), doi:10.1016/j.jnca.2010.04.004
4. Deruyck, M., Joseph, W., Lannoo, B., Colle, D., Martens, L.: Designing Energy-Efficient Wireless Access Networks: LTE and LTE-Advanced. IEEE Internet Computing 17(5), 39–45 (2013), doi:10.1109/MIC.2013.6
5. Deruyck, M., Vereecken, W., Joseph, W., Lannoo, B., Pickavet, M., Martens, L.: Reducing the power consumption in wireless access networks: overview and recommendations. Progress in Electromagnetics Research-PIER 132, 255–274 (2012)

6. Deruyck, M., Vereecken, W., Joseph, W., Lannoo, B., Pickavet, M., Martens, L.: Reducing the power consumption in wireless access networks: overview and recommendations. Progress in Electromagnetics Research 132, 255–274 (2012), doi:10.2528/PIER12061301
7. Erceg, V., Schumacher, L., et al.: IEEE P802. 11 Wireless LANs. TGn Channel Models, doc.: IEEE, pp. 802–11
8. Joseph, W., Verloock, L., Goeminne, F., Vermeeren, G., Martens, L.: Assessment of general public exposure to LTE and RF sources present in an urban environment. Bioelectromagnetics 31, 576–579 (2010)
9. Jourdan, D., de Weck, O.: Layout optimization for a wireless sensor network using a multi-objective genetic algorithm. In: IEEE 59th Vehicular Technology Conference, VTC 2004, vol. 5, pp. 2466–2470 (2004), doi:10.1109/VETECS.2004.1391366
10. Koutitas, G., Samaras, T.: Exposure minimization in indoor wireless networks. IEEE Antennas and Wireless Propagation Letters 9, 199–202 (2010), doi:10.1109/LAWP.2010.2045870
11. Lee, J.W., Choi, B.S., Lee, J.J.: Energy-efficient coverage of wireless sensor networks using ant colony optimization with three types of pheromones. IEEE Transactions on Industrial Informatics 7(3), 419–427 (2011), doi:10.1109/TII.2011.2158836
12. Liu, N., Plets, D., Goudos, S., Joseph, W., Martens, L.: Multi-objective network planning optimization algorithm: Human exposure, power consumption, cost, and coverage. Wireless Communications and Mobile Computing (submitted)
13. Nagy, L.: Indoor radio coverage optimization for WLAN. In: 2nd European Conference on Antennas and Propagation (EuCAP 2007), p. 225 (2007), doi:10.1049/ic.2007.1348
14. Plets, D., Joseph, W., Vanhecke, K., Martens, L.: A heuristic tool for exposure reduction in indoor wireless networks. In: 2012 IEEE Antennas and Propagation Society International Symposium (APSURSI), pp. 1–2 (2012), doi:10.1109/APS.2012.6348505
15. Plets, D., Joseph, W., Vanhecke, K., Tanghe, E., Martens, L.: Development of an accurate tool for path loss and coverage prediction in indoor environments. In: 2010 Proceedings of the Fourth European Conference on Antennas and Propagation (EuCAP), pp. 1–5 (2010)
16. Plets, D., Joseph, W., Vanhecke, K., Tanghe, E., Martens, L.: Coverage prediction and optimization algorithms for indoor environments. EURASIP Journal on Wireless Communications and Networking, Special Issue on Radio Propagation, Channel Modeling, and Wireless, Channel Simulation Tools for Heterogeneous Networking Evaluation 1 (2012)
17. Plets, D., Joseph, W., Vanhecke, K., Tanghe, E., Martens, L.: Simple indoor path loss prediction algorithm and validation in living lab setting. Wireless Personal Communications 68(3), 535–552 (2013), doi:10.1007/s11277-011-0467-4
18. Ran, M., Ezra, Y.B.: Green femtocell based on uwb technologies. Novel Applications of the UWB Technologies, 175–194 (August 2011)
19. Saunders, S.R.: Antennas and Propagation for Wireless Communication Systems. John Wiley & Sons Ltd. (1999)
20. Unger, P., Schack, M., Kurner, T.: Minimizing the Electromagnetic Exposure Using Hybrid (DVB-H/UMTS) Networks. IEEE Transactions on Broadcasting 53(1), 418–424 (2007), doi:10.1109/TBC.2006.889207
21. Verloock, L., Joseph, W., Vermeeren, G., Martens, L.: Procedure for assessment of general public exposure from wlan in offices and in wireless sensor network testbed. Health Physics 98, 628–638 (2010)

22. Vilovic, I., Burum, N., Sipus, Z.: Design of an indoor wireless network with neural prediction model. In: The Second European Conference on Antennas and Propagation, EuCAP 2007, pp. 1–5 (2007)
23. Vilovic, I., Burum, N., Sipus, Z.: Ant colony approach in optimization of base station position. In: 3rd European Conference on Antennas and Propagation, EuCAP 2009, pp. 2882–2886 (2009)
24. Yun, Z., Lim, S., Iskander, M.: An integrated method of ray tracing and genetic algorithm for optimizing coverage in indoor wireless networks. IEEE Antennas and Wireless Propagation Letters 7, 145–148 (2008), doi:10.1109/LAWP.2008.919358

A Context-Aware Framework for Media Recommendation on Smartphones

Abayomi M. Otebolaku and Maria T. Andrade

Abstract. The incredible appeals of smartphones and the unprecedented progress in the development of mobile and wireless networks in recent years have enabled ubiquitous availability of myriad media contents. Consequently, it has become problematic for mobile users to find relevant media items. However, context awareness has been proposed as a means to help mobile users find relevant media items anywhere and at any time. The contribution of this paper is the presentation of a context-aware media recommendation framework for smart devices (CAMR). CAMR supports the integration of context sensing, recognition, and inference, using classification algorithms, an ontology-based context model and user preferences to provide contextually relevant media items to smart device users. This paper describes CAMR and its components, and demonstrates its use to develop a context-aware mobile movie recommendation on Android smart devices. Experimental evaluations of the framework, via an experimental context-aware mobile recommendation application, confirm that the framework is effective, and that its power consumption is within acceptable range.

1 Introduction

The advancements in mobile computing technologies, wireless and mobile networks, and the proliferation of mobile devices such as smartphones have brought remarkable changes in the way we access online media items. Using smartphones via the Internet to access media content has become easier and ubiquitous. Moreover, with mobile phones now equipped with sophisticated video cameras and multimedia functionalities, user generated media content has become pervasive. Because of this development, much more media content is published every moment online. Consequently, mobile users can obtain myriad choices of

Abayomi M. Otebolaku · Maria T. Andrade
Telecommunications and Multimedia Unit, INESC TEC, Portugal
e-mail: abayomi.otebolaku@inescporto.pt

Maria T. Andrade
Faculty of Engineering, University of Porto, Portugal
e-mail: mandrade@fe.up.pt

L. De Strycker (ed.), *ECUMICT 2014*,
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_8, © Springer International Publishing Switzerland 2014

media items to consume. Nevertheless, this development comes with negative effects, owing to the exponential explosion of online-based media contents. Because of the huge volume of online-based media content, mobile users now waste valuable time hoping to find relevant ones to consume. Often, because mobile user's preferences are subject to contextual situations, they end up with irrelevant media items, which do not match their preferences [2-5, 7-8, 10-11].

Therefore, to assist mobile users make informed choices about relevant content, many works have been developed [1], [15]. The highly successful traditional approaches, based on personalized recommendations assist users to find relevant items by using evaluations of the previously consumed contents given by the target user (content based) or the evaluations given by users who are similar to him (collaborative based). These traditional methods fall short of getting relevant media items to users because they assume that users always give an evaluation of content they consume, which in practice rarely happens. Additionally, they do not consider contextual information as an important factor that affects user's preferences [2, 18].

Recently, however, context information has become integral part of the recommendation process. For example, context information is used to generate music recommendations for mobile users in [11]. In [12], context information is explored to recommend interesting movies to mobile users. However, these systems use context explicitly to recommend media content to users, thereby requiring user's constant interventions. Users would like the system to implicitly suggest content that match their preferences, without their interventions.

To deliver this rich personalized media experience to mobile users, and to address the weaknesses of the traditional approaches, especially collaborative and content based approaches [21], it is important to understand the relationship between a user's preference, context information, and media services. This relationship influences the media content consumption choices of mobile users.

As illustrated in Figure 1, the contribution of this paper is the provision of a generic framework with features that can automatically suggest content to users, based on their dynamic contexts, inferred from their smartphone-embedded sensors. The main functionality of the framework is the capturing and the use of context information, especially the mobile user activities, which it relates to the user preferences to select suitable content among available online media, personalizing it for specific users and their contexts. Furthermore, we have implemented the framework's components as RESTful web services [22], making it accessible to any platform, and we have evaluated its functionality and feasibility.

In the next section, we analyze briefly other research efforts for media item recommendations, and then describe the problems that the framework addresses in section 3. Section 4 presents the framework's conceptual design, and section 5 describes its implementation and some evaluations. Section 6 discusses the evaluation results. Section 7 concludes the paper and gives our future direction.

2 Related Work

Most popular work in context aware recommendation such as Pessemier et. al. [3], Chen [5], and Yu et.al. [10], focus on explicitly on using context information to

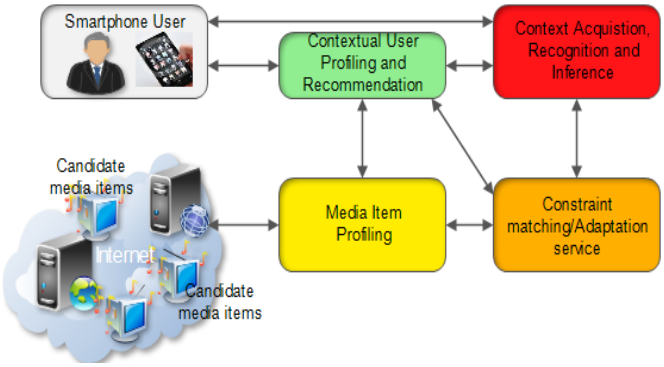


Fig. 1 Proposed framework's functional view

suggest media items. Cinemappy [12] uses the location and movie consumption history of mobile users to recommend relevant movies. Our work integrates a context model that accommodates broader contextual information than just location information. In addition to user location, our work integrates the user as well as user environment information. Whereas, [11] uses user activities such as walking, running, studying, etc. only to recommend music to mobile users, its weakness is that it considers only user activity and not location information or other context information that influence user preferences.

Our approach is different because the framework proposes a hybrid and generic system that provides both explicit and implicit media recommendations to mobile users. The context recognition process is exposed as a web service to be consumed by any application built on top of the framework. The contexts can be stored as historical data, which can be related to the present user contexts and the user preferences. Similarly, the framework exposes its recommendation processes, and contextual user profiles as web services, providing easy access from any mobile software and hardware platforms. It also includes an optional adaptation service that can tailor the presentations of the recommended contents to the device and network constraints.

3 Problem Definition and System Requirements

Enabling proactive contextual suggestion and delivery of relevant rich media items that satisfy the mobile user's preferences, in heterogeneous environments requires careful design considerations, because providing media items to users in such an environment is a challenging problem. This requires addressing several key issues, which are dynamic context and activity recognition, representation, contextual user profiling, preference management, an adaptation of the media item

presentation to device constraints and network prevailing conditions, etc. Besides, developing a framework that can incorporate these aspects in a unifying framework has never been trivial. In this section, we examine the difficulties that characterize the design of the proposed framework.

3.1 Context Recognition and Representation

An item is relevant if it belongs to the same context as the user's active interaction [14]. Therefore, wrong context will lead to wrong recommendation. To explore contextual information for personalized content delivery, it is important to identify accurately, the user's contextual situations, which require context sensing capabilities. However, sensed data are generally vague and imprecise, because they contain noise, and in raw form, they do not make sense [18]. Additionally, several sensory data need to fuse to recognize more meaningful high-level contexts. These conditions necessitate a context recognition model, utilizing machine-learning approach that takes the low-level sensory signals as inputs and produces as output, accurate and meaningful high-level context information. Furthermore, recognized high-level context information must relate to other high-level context information to infer more semantically meaningful contextual situation of the user. The dynamic realization of this requirement remains a challenge in the overall process. Section 4.1 discusses how this issue is addressed by CAMR.

3.2 User and Media Content Profiling

A key requirement to offer truly contextual delivery of media content to mobile user is the user profile. The user profile should encode all desirable media content features, customized for the profile owner and in context. The user profile includes, among others, preferences of the user for 1) desired media presentation characteristics, (2) optional user identities such as names, gender and profession, 3) preferred location information, 4) usage history and, 5) other high-level context information such as user activities in relation to the user's preferences. Dynamic acquisition of user preference information, user consumption history, and their incorporation with user's dynamic context and activities to update the user's changing preferences and estimate their relevance and importance remains a challenge. Section 4.2 discusses the user profiling issue in CAMR.

3.3 Content Classification and Recommendation

The Web today contains a massive amount of digital contents that users explore using their mobile devices. The sheer size and the decentralization of these digital content make it difficult for mobile users to obtain those contents that suit their preferences and situations. Recommendation algorithms provide the advantage of

guiding users in a personalized way to interesting content in a large space of possible options [1]. However, the recommended content must be provided according to the user's changing contexts, activities, and preferences. Therefore, content classification and recommendation need to be equipped with contextual information to assist mobile users to get rich media experience while on the move. Section 4.3 presents the classification and recommendation process in CAMR.

4 Context Aware Framework for Media Recommendation

In this section, we present CAMR as a conceptual framework for generating contextually relevant media content to mobile users. CAMR addresses the problems discussed in the last section. Figure 2 sketches the architecture of the framework, showing its major components. These components are described in this section.

4.1 Context Recognition

The ability of a system to identify contextual situations and respond to them is one of the most important functions of any context-aware system. However, context sensors emit data that are in low-level form, which are not suitable for mobile application. Context recognition is a process that collects raw data from sensors and transforms them to information that can be used by applications [18]. To provide dynamic contextual information about media items consumers, CAMR uses context recognition process to identify contextual information such as user activities, user location, and environment situation such as weather, illumination, and noise level from low-level data collected from smartphone-embedded sensors. To realize this, the context recognition service leverages four important processes. The next section presents these processes.

Sensor Data Collection and Preprocessing

CAMR gathers events from smartphone embedded sensors such as accelerometer, GPS, gyroscope, rotation vector, orientation, proximity, microphone and light sensors. It collects 128 samples of data from each axis of accelerometer, rotation vector, and orientation sensors in a continuous 3 seconds, with 64 samples from the previous 3 seconds overlapping the next 3 seconds. In the data preprocessing phase, we removed event outliers [9, 13]. This is done by removing samples from the beginning and the ending of each example to reduce the influence of noise in the data.

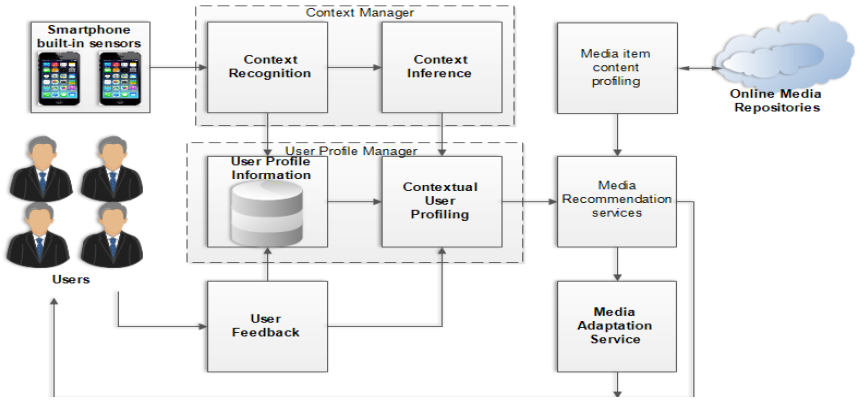


Fig. 2 CAMR architecture

Feature Extraction

User's dynamic contexts, such as activities, are performed in relatively lengthy period, in seconds or minutes, when compared with the sampling rate of the sensors. The sampling rates usually do not provide sufficient data to describe user activities. Therefore, activities are usually identified in the time window basis rather than sampling rate basis [13]. Comparing a series of time windows to identify activities is almost impracticable even if the signals being compared come from the same user performing the same activity context [18]. The feature extraction process addresses this problem by filtering relevant and obtaining quantitative data from each time window. Various approaches have been explored in literatures such as statistical properties and structural properties of the sensor signals. Structural features such as Fourier transform are quite complex and require more computational resources. This may not be ideal for resource starved devices. On the other hand, statistical features are simple and require less computational resources [9]. Thus, CAMR uses simple labeled statistical features [range, maximum, minimum, mean and standard deviation], which have been validated in our previous work to be very effective, to discriminate between time windows [9]. These features are extracted into feature vectors that are then used in the next process.

Context Classification

After extracting the time window features from the raw sensor signals, without deriving the context knowledge from them, the example features are meaningless. CAMR uses classification algorithms, in particular, Support Vector Machine, Neural Network (NN), Decision Trees, Nearest Neighbors (KNN), and BayesNet to derive high-level context from the statistical features. Details of the modeling and evaluations can be found in [9, 13, 18]. The present implementation integrates KNN in the recognition service to accurately obtain independent future activities and contexts of the mobile users. We implemented KNN based context

recognition model because evaluations of various models in our previous report [9] confirmed its excellent performance in terms of accuracy and recognition time.

Context Inference

The contextual information obtained from the preceding process is usually semantically not adequate for recommendation process. For example, it is important to know what a user is doing, when, and where he is doing it. Nevertheless, without relating this information to provide situational context, this information will convey no useful sense. This is one of the weaknesses of the existing work. CAMR uses knowledge-based model on top of the classification process to relate different atomic context information to obtain contextual information at a higher semantic level. For example, having known that a user sitting at home is located in the living room, if we know that the TV is switched on (can be obtained from IR Blaster), and then one can relate the two information and conclude that the user is watching TV. The user is watching TV, obtained from sitting, home, living room, TV, etc. is inferred using the ontology based knowledge inference process of CAMR. The details of this knowledge based process have been presented in [6], and will not be discussed further in this article. CAMR can also determine such complex context, such as a user is "*jogging in the sport arena*". This situational information is crucial to offer a rich media experience to mobile users.

4.2 Contextual User Profiling Service

A user profile describes his preferences, normally based on the history of the user's actions [1]. CAMR's contextual user profile service (CUPS) summarizes the user's content consumptions into a limited set of categories. Categories are characterized by one or more genre, and a number of properties characterize the genre. Several genres can be associated to one category. Several properties can be associated with one genre. Additionally, it incorporates the contextual dimension, associating one or more inferred context to each *category-genre-property* concept. It presents each user profile as a four level tree, as shown in Figure 3 (a), with the root of the tree representing the user's optional demographic information. The first level of nodes corresponds to the category; the second level represents the genre; the third level contains the properties of a given *category-genre*. This level provides the media item's context, characterizing at a finer detail, the consumed content and thus the preferences of the user. A limited set of properties is used for each genre to obtain a good compromise between sufficient degree of characterization of content (hence, sufficient ability to make distinctions) and reasonable dimensions of the user profile. The leaf nodes provide information about the contexts where the user preferences have been observed. Leaf nodes have three fields – *type*, *intensity,* and *lifetime* – whereas all other nodes have only the type field. In the leaves, the types represent the type of context. The newly

introduced concepts in the user profile, the intensity and lifetime track user's contextual consumption history.

Using these weighted parameters, the system is able to determine at runtime, the media contents that are important to the user, based on his contextual preferences. The *intensity* provides information on the number of times the user has consumed items of that *category-genre-property* in that specific context. The intensity (the dynamic preference of the user) of those elements belonging to the media's term is obtained by summing up the products [weight x lifetime] of all their occurrences. The intensity value of the retained elements at that level is obtained by visiting their child nodes in a breadth-first traversal. The same applies to the retained elements of the category level. The intensity of the elements belonging to the genre level is the largest value of their children. Accordingly, these values are obtained by performing a depth-first traversal. This way, the user profile can handle any category of media items such as movie, news, music, etc.

To classify the candidate media items, and to obtain a list of recommended items, a vector is created from the user profile. A global version of the user vector contains as many elements as the number of different *category-genre-properties* that appear in the complete user profile tree.

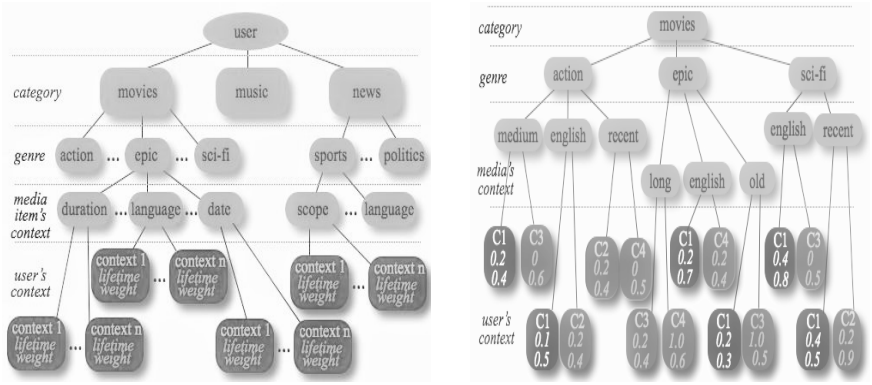


Fig. 3(a&b) Profile Structure for a hypothetical user

A *contextualized user vector* typically has a much smaller number of elements, corresponding to the different category-genre-properties associated with the specific context under consideration. The *contextualized user vector,* V_c is, thus, built using context data to filter the profile. The value of the current context of usage is compared to the leaves of the profile tree (context nodes) to identify the upper nodes that provide values for the elements of V_c. Only nodes whose leaves match the current context are used. Each element in the vector is a pair *keyword-intensity*. *Keyword* is the textual value of nodes (the *type* field); *intensity* is obtained by multiplying the values of the fields' *weight* and *lifetime* of the respective node. As an example, consider the hypothetical user profile of Mr. X represented in Figure 3 (b). Assuming that the system has inferred that Mr. X is in context C_1, the elements that will be included in the contextualized user profile

vector are the ones that have leaves with context value C_1. The intensity of those elements belonging to the media's context is calculated by summing up the products [*weight* x *lifetime*] of all their occurrences (e.g., the node with value "English" occurs three times for context C_1; therefore the intensity of "English" is the sum of the corresponding three intensities: 0.1 x 0.5 + 0.2 x 0.7 + 0.4 x 0.8 = 0.51). The intensity value of the retained elements at this level is obtained by visiting their child nodes, using a breadth-first traversal. The same applies for the retained elements of the category level. The intensity of the elements belonging to the genre level is the largest value of their children. Accordingly, these values are obtained, using a depth-first traversal.

4.3 Contextual User Profile Information Store

The contextual user profile information store (CUPIS) is the persistence store where user profile data such as lifetime and intensity of each user preference, contexts, etc. are kept. Consequently, interactions of the user with the systems, including his feedback are likewise stored for subsequent recommendation processes. User feedback manager (UFM) and CUPS rely on CUPIS for persistence.

4.4 User Feedback Manager

It is important to track user's consumption behavior to improve the system's classification accuracy. The user profile can be updated using two approaches, explicit and implicit methods [1]. The former grants the users the ability to modify the values assigned to their preferences by the system. The implicit approach involves preference learning without direct user intervention, such as updating the profile when the user has spent a certain amount of time on a given item. The CAMR's user feedback manager (UFM) updates the user profile whenever the user interacts with the system; it tracks both consumption and non-consumption of content by the user to learn the contextual preferences of the user for any kind of media item. To implicitly learn the user preferences, UFM runs an intermittent background service monitoring the interactions of the users. Additionally, it explicitly updates the profile whenever a user consumes any content item by obtaining a feedback from such user. Equations *(1&2)* are used for the implicit user profile update model, allowing associating importance to the finer details of content and corresponding contexts of consumption. The model associates weight $w_i m_i \in [0,1]$ that relates the relevance of a consumed media content to the context in which it is consumed, learning the user preference for the content he consumes or those he does not consume. A weight w_{ij} represents the relevance of content m_j belonging to the media content category k_i that a user u_i consumes in context c_i. Whenever the recognition model detects that the user is in a context c_i, for a continuous period of time [0, T] and the user consumes one or more of content m_1, m_2,…m_n recommended by the recommendation service, then weight $w_i m_i \in [0,1]$ is associated with this content, which at time T is updated as follows:

$$w_i m_i = w_i m_i + \gamma(\alpha - w_i m_i)\ i = 1,2,3,\ldots,n. \quad (1)$$

Then for those content b_1, b_2, $b_3,\ldots,b_n$ with associated weights w_1b_1, w_2b_2, $w_3b_3,\ldots,w_nb_n$ not consumed by the user, these weights are updated as follows:

$$w_i m_i = w_i m_i + \gamma(\alpha - w_i m_i)\ i = 1,2,3,\ldots,n\ . \quad (2)$$

γ is a learning parameter whose value is obtained by: $\alpha \in [0,1]$ its value is set to 1 in (4) and 0 in (5)

Factor *t* in equation (3) represents the number of days elapsed since the last time the user has consumed an item with the characteristics described by his profile nodes. With equation (3), the parameter γ for each node remains above 0.9 during the first 30 days after it has been visited, rapidly decreasing to zero after that period (non-negative values are automatically converted to zero). For all other nodes, its value is linearly increased daily. This way, nodes that represent items that have not been seen for a long period, will eventually have no impact on the user preferences evaluation.

$$\gamma = 1 - \left(\frac{t}{45} \right)^5 \quad (3)$$

4.5 Media Item Content Profile Service

The media item content profile service (MICPS) is responsible for crawling the Internet for candidate media items. It retrieves and processes the media item metadata into a form that can be processed by the recommendation service. Usually, it filters the metadata and scores the terms in the metadata to generate a media item profile vector, which is fed into the recommendation services. It is thus necessary to create a media vector, V_M, for each media item. To describe the media items, MICPS relies on the availability of semantic metadata using the MPEG-7 MDS semantic tools [20]. Given that in practice most of the media resources available online lack this metadata, our system incorporates alternative methods to obtain the semantic descriptions of the candidate content. One of such alternatives is the Internet Movie Data Base (IMDB) service API, Last.fm API etc. For each media item, a vector V_M is initially created as an exact replica of V_C. Then, for every element of V_M, the system inspects the MPEG-7 metadata for a match. If it finds a match, it retains the intensity value of the matching element in a V_M. Otherwise, it assigns zero to the element.

4.6 Media Item Recommendation Service

The recommendation service explores three recommendation algorithms for context-aware media recommendations. The content base (CBF), the collaborative based (CF) and a hybrid based recommendation algorithms. In this article, we will only elaborate on the hybrid approach, which is based on context-aware content-based collaborative process. The traditional collaborative recommendation

generates predictions for the target user based on the item previously rated or viewed by other users, and the content-based approach, which uses the consumption history of the user. These two approaches suffer from the so-called overspecialization and new user/item problem respectively, which excludes casual users or those whom the system does not have enough information to generate recommendations [1,16], or always suggesting the items similar to those consumed in the past by the user. Hybrid recommenders combine one or more of the conventional recommendation processes to overcome their individual weaknesses to gain better performance [1]. To take advantage of the hybrid technique, we built context-aware content based collaborative recommendation (CACBR), a hybrid recommendation that combines the context-aware CF and CBF. Basically, it uses the contexts in which other users have consumed the content previously to find users that are similar to the target user by comparing the active user context history and the target user's present context. This way, CACBR can address the new user/item problem of collaborative process and the overspecialization problem of the content based process. This is achieved in three phases. In the first phase, it identifies every user (neighbor) that is similar to the target user by searching through each user's profile tree, looking for context that matches the target user's recognized context. For every user profile with a match, the intensity value p_{ni} of the *category-genre-property* nodes in the user profiles is retrieved into a vector. The vector is then used to calculate the similarities, using equation (5), between all users. After this calculation, it then selects the *top n* most similar users, called neighbors or friends of the target user who have consumed content in the same or similar contexts to the target user's current context. In the second phase, it ranks the candidate content for each neighbor by obtaining vectors V_c and V_m corresponding to contextual user profile of every neighbor and candidate media profile vectors respectively.

$$Sim\ (v_c, v_m) = \frac{V_c . V_m}{V_c \times V_m} = \frac{\sum_{i=1}^{n} V_{c_i} \times V_{m_i}}{\sqrt{\sum_{i=1}^{n} (V_{c_i})^2} \times \sqrt{\sum_{i=1}^{n} (V_{m_i})^2}} \tag{4}$$

By applying the cosine formula (4), it calculates the distance between the contextual user profile vector (CUPV) V_c and the media item profile vector (MIFV) V_m.

$$Sim(P_u, P_n) = \frac{\sum_{i \subset CP_{u,n}} (p_{u_i} - \overline{p}_u)(P_{ni} - \overline{p}_n)}{\sqrt{\sum_{i \subset CP_{u,n}} (p_{u_i} - \overline{p}_u)^2} \times \sqrt{\sum_{i \subset CP_{u,n}} (p_{n_i} - \overline{p}_u)^2}} \tag{5}$$

In the third phase, it generates the preference prediction value for each of the content from (5) for the target user, using Resnick [17] prediction formula (6). In this formula, C_i is the intensity to be predicted for each candidate content *i* generated from formula (5) for the target user and p_{ni} is the intensity by the

1) *Sense and classify target user present context*
2) *Search target user's profile to find if current context matches any context in his profile*
   a. *If it matches:*
      i. *Generate his contextual profile vector*
      ii. *Generate contextual profile vector for every other user in the profile repository*
      iii. *Generate all users similar to the target user(using the above contextual profile vector and Pearson correlation) based on equation(5)*
      iv. *Retrieve media metadata from the Internet*
      v. *Generate media content profile vectors*
      vi. *For every similar user in(iii),generate its similarity value with every media content obtained using (4)*
      vii. *Rank the candidate contents for each similar user*
      viii. *Generate a preference prediction for target user for every media item using modified Resnick formula(6)*
      ix. *Rank the TopN candidate content from (vii) for the target User*
   b. *If no match is found:*
      i. *Generate non-contextual user profile vector*
      ii. *Generate non-contextual user profile vector for every user in the profile repository*
      iii. *Generate all friends of the active user(using the non-contextual user profile vectors)*
   iv *Repeat a(iv-x) above*
   c. *Present suggested items to target user*
   d. *record feedback(implicit/explicit)*
   e. *Log the present context and update his user profile*

Fig. 4 Context aware content based collaborative process

$$C_i = \overline{C} + \frac{\sum_{i \subset neighbors(u)} (p_{n_i} - \overline{p}_n) Sim(c, p)}{\sum_{i \subset neighbors(u)} |sim(c, p)|} \quad (6)$$

$\overline{C}$ is the average intensity of all terms in target user's profile, whereas $\overline{p}_n$ is the average intensity of all terms in neighbor *i* profile. The *sim (c, p)* is the similarity measure between profiles of the target user *c* and neighbor *p*, calculated using formula (4). The modified Resnick prediction discounts the contribution of every neighbor's intensity according to its degree of similarity with the target user so that more neighbors have a large impact on the final intensity predictions [17].

Finally, recommendation list is built by ordering the candidate media items in descending order of magnitude of the computed prediction values. Figure 4 summarizes the entire hybrid content classification (CACBR) process.

5 Framework Implementation and Evaluation

This section presents the implementation of CAMR and its evaluation via a mobile movie recommendation application that was built on top of it. First, we present the framework integration, second its implementation and third, its evaluation.

5.1 Framework Component Integration

We describe, in this section, the integration of different aspects of CAFMR. Figure 5 shows the component model of CAMR. It provides a detailed version of the high-level architecture shown in Figure 2. In this figure, all the components

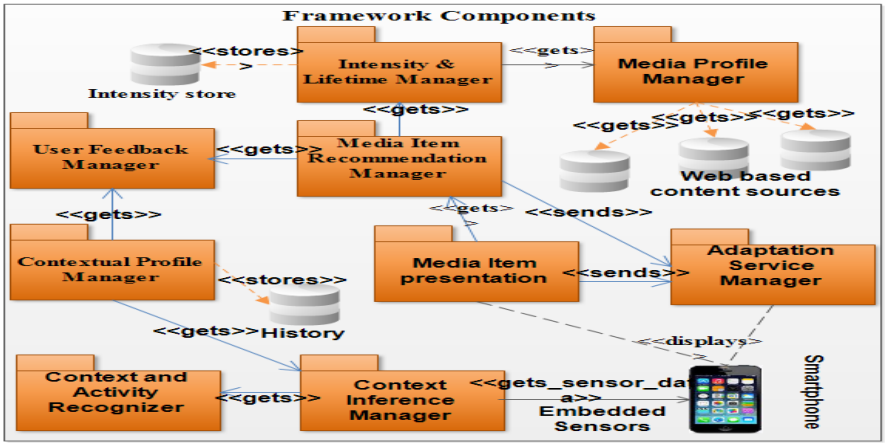


Fig. 5 CAMR Component Model

interact via interfaces. *Media Item Recommendation Manager* is the central component of the framework, establishing direct or indirect connections between the other components. It provides the service that executes the task that generates the recommendations. The *Media Profile Manager* connects the online-based databases such as YouTube, IMDB, etc. to extract media metadata descriptors. The *Context and Activity Recognizer* connects with various sensors on the smartphone to obtain raw sensor data, which it then processes to obtain high-level context information. The *Context Inference Manager* is responsible for relating two or more recognized context information to derive semantically expressive high-level context information. The *Contextual User Profile Manager* collaborates with *Context Inference Manager* and *Context and Activity Recognizer*, taking the context information, user preferences, and establishing a relation between the context information and user preferences.

The *User feedback manager* is responsible for learning the user's interaction with the framework; it collects either explicitly or implicitly, the user's contextual feedback in order to improve the future recommendations, which truly reflects the user's contextual preferences. *The intensity and lifetime manager* is responsible for scoring the contextual preferences of the user, and for tracking those media items that users have consumed before, which are no longer interesting to them. The *Media presentation manager* is responsible for determining the appropriate format of the recommended media items to display by the smartphone. The *adaptation manager* is responsible for determining if the presentation format of the recommended media items can be played by the smartphone. If it cannot be played, it then determines the appropriate adaptation mechanism to be executed to display the media item.

5.2 CAMR Implementation

To demonstrate the feasibility of CAMR, we have implemented its prototype consisting of eight services, representing the operations offered by its components. The context recognition component is implemented as a mobile client service that can be deployed on smartphones. Other components of Figure 5, including the media item recommendation and the contextual user profile managers have been implemented using Java technologies, integrating Java EE 7 (EJB, JPA) and RESTful Web Services and deployed on the Oracle Glassfish server [19]. MySQL 5.6 database [22] serves as the backend database, hosting user profiles and user context history. The contextual user profile information service registers user actions and performs all the necessary processing such as the user profiling,

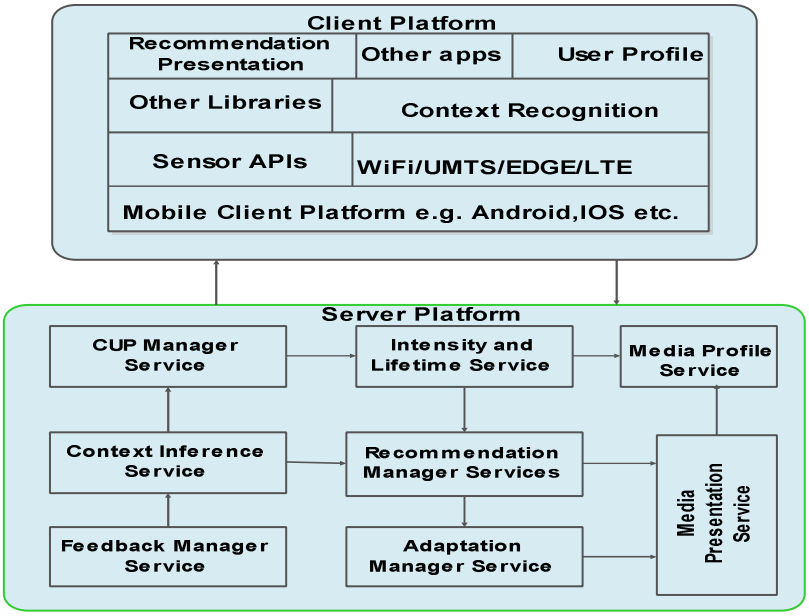


Fig. 6 Implementation Architecture

update, whereas the recommendation generation is performed by media recommendation service, at the same time working in tandem with the context service and recommendation client service on the Android device. The client-server implementation architecture is depicted in Figure 6. Figures 7&8 presents some screenshots of the mobile client. Figure 7 (a) shows the user profile management interface, where a user can optionally manage her profile. In the application, three categories of contents (movies, music, and news) are presently supported. Figure 7 (b) is the context browser where contextual information can be managed on the device by the users. It shows the high-level context indicating, for example, that the user is at (*Home)* location. The user's activity is inferred as *Sitting* while she is *indoors,* and time of the day inferred as *night.* This context

information is fed into the context knowledge base, which infers higher level situational context, such as "*It is weekend night, user is sitting at home*". This is then fed into the contextual user-profiling model, which is stored in a context history repository. The process for deriving high-level context information is based on context recognition using context classification algorithms described in section 4. The figure also shows that the user is indoors, which has been obtained using the device built-in GPS. Figure 8 (a&b) shows the recommendation interfaces of the context-aware mobile recommendation application, showing the recommended movie content and all options available to user to visualize the recommendation such as playing the movie trailers, or connecting to online sources for additional information on the recommended items.

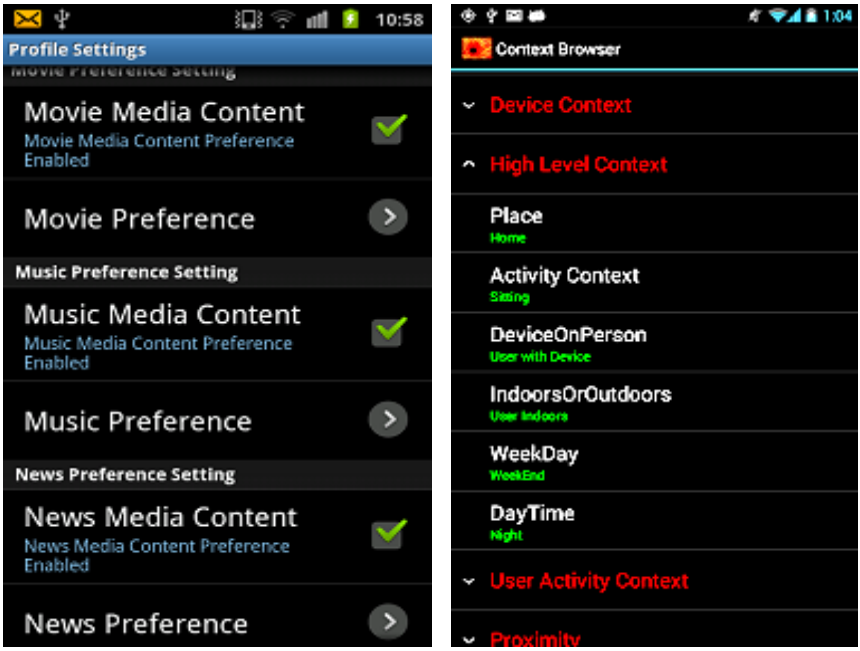


Fig. 7 (a) content preferences. (b) context browser.

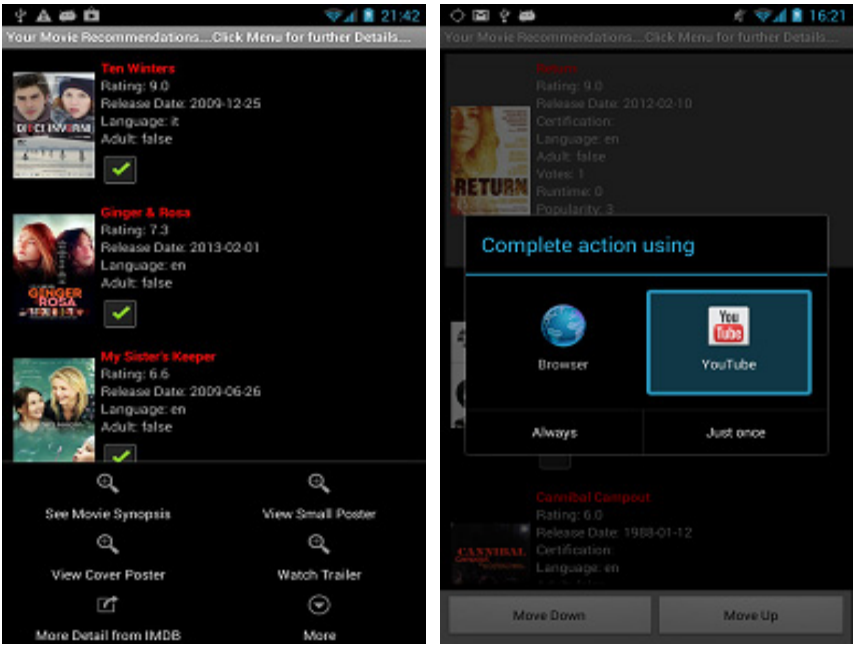


Fig. 8 (a) recommendation list (b) recommendation presentation

5.3 Experimentation

To experiment and to evaluate the feasibility of the proposed CAMR, we have implemented the framework, and based on it developed a context-aware movie recommendation shown in Figures 8(a&b). In this section, we present the evaluation of the framework.

The Experimental Data

To evaluate the feasibility of the proposed framework, preliminary experiments have been conducted using two sets of data. First, the candidate dataset was obtained by crawling over 4500 movie metadata records from the Movie Database (themoviedb.org), further enhanced with additional metadata retrieved from the IMDB.com. This metadata set contains 23 different movie genres, and each record contains on average three different genre labels. Language, cast, country, duration, and release date characterize the genres. These terms, thus constituted the media item's context in our user profile model (as illustrated in Figure2). Second, we solicited real world user profile data from 200 online users, each having 19 different entries in the genre level (the entry in the category level was the same for all users – movies). High-level contexts such as Location = "*Home*", DaysOfWeek ="*WeekEnd*", TimeOfDay = "*Night*", Activity ="*Sitting*" were associated with these entries.

The Evaluation Metric

We need to define a way to measure the impact of contextual information on the recommendation quality. We do this by comparing the recommendations with or without activity contexts. We define a recommendation quality metric (also known as precision), which is the percentage of the number of times CAMR successfully provides mobile users with preferred media items. A provided recommendation is relevant if the user finds what he wants within the first N provided recommendation list. For example, if the user finds n^{th} media item of a given recommendation list, then the recommendation is successful if and only if $n \leq N$. Having a higher accuracy with lower value of N is low means that the recommendation framework works well and that its contextual user profile accurately presents the user's preferences.

Experiments

This section presents the evaluation of the framework; we evaluate the accuracy of the context recognition service, the quality of the recommendation and the energy consumption of the framework.

A) CAMR recommendation quality

Experimental Procedure: To evaluate the quality of the recommendations, given that most of the 200 users were anonymous and thus were not available to provide continuous on-device feedback, we devised an approach to allow marking recommended items as *relevant* or as *irrelevant*. This allowed us to simulate the acceptance or rejection of suggested content by the users as shown in Figure 8 (a). In the evaluation process, an item is marked as relevant provided that at least 2/3 of the terms that appear in its metadata record (but not less than two terms), also appear in the user profile with an intensity larger than the average intensity of all terms in the user profile. The position (n) of the item in the recommendation list of N items must fulfill the condition that $n \leq N$. For example, a suggested movie item with a metadata record presenting three terms is marked as relevant if two out of those three terms matching the user profile, with intensities larger than the average intensity of all terms. Otherwise, it is marked as an irrelevant media item. We adopted this approach because we observed in our experiments that the classification of candidate items is also influenced by the number of terms contained in their metadata records, particularly in the *genre-property* nodes. We defined two scenarios to experiment and to evaluate the quality of the recommendation service.

1) The first situation involves generating recommendations based on *category-genre-property* nodes in the user profile. This scenario is realized by using content-based recommendation and collaborative based filtering as the recommendation algorithms, without contextual information. *(ii)* The second scenario is similar to the first. However, in this experiment the recommendation is based on the *category-genre-property-context* nodes of the user profile tree. In other words, this is a situation where contexts play very significant role in the user's preferences i.e. contextual recommendation where the system generates recommendations using

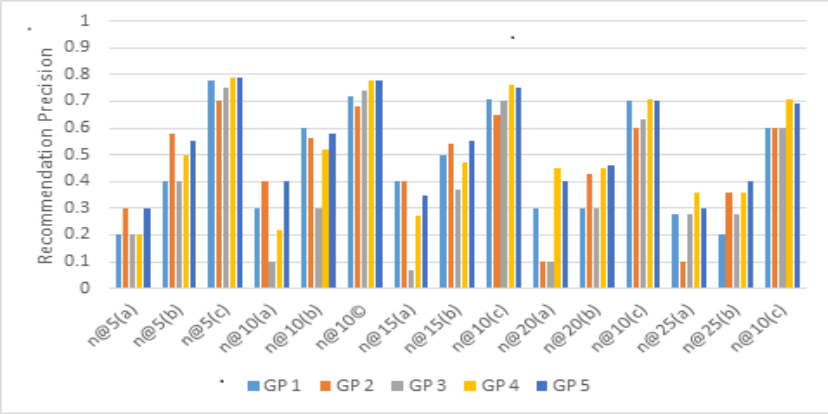


Fig. 9 Contextual recommendation prediction (Precision)

user's contextual information, especially user's activities. The 200 user profiles are divided into 5 by grouping them according to their similarities. We then issued recommendations based on the context-aware hybrid algorithm and the traditional algorithms to a user, representative of each group. The recommendation is repeated for n = *5,10,15,20,* and *25*. Figure 9 is the precision plot showing the recommendation quality for each group, where a, b, and c represents the precisions for CF, CBF, and the CACBR.

B) Energy consumption evaluation: Since the proposed solution involves sensor monitoring to recognize user contexts and activities, battery drain becomes a concern for the smartphone users. GPS, WiFi or cellular network, accelerometer, orientation, rotation, and other sensors have a significant impact on the battery lifetime. Therefore, to evaluate the impact of the context recognition service on the battery lifetime, the context recognition service was deployed on a Samsung Galaxy S I9000 smartphone as the only active application. We then evaluated the system in eight situations, using the power tutor application [23]. These situations and the percentages of power consumed are shown in table 1. The table shows that the smartphone screen consumes a lot of energy in the active state.

Table 1 Context recognition power consumption

	Sensors	Percentage of Power Consumed
1	All sensors+ Screen	40.66
2	All sensors - Screen	20.84
3	Periodic 30 s + GPS+WiFi+all	8.90
4	Periodic 60 s + GPS+WiFi+all	7.35
5	GPS + WiFi+Accelerometer	5.60
6	GPS+WiFi+Orientation+other	10.78
7	GPS+WiFi+rotation+other	10.40
8	Wi-Fi +other sensors-GPS	7.55
9	GPS +other sensors -WiFi	8.5

Therefore, the context recognition service has been implemented as a background service running at predetermined intervals. This is an important energy optimization consideration in our implementation. In addition, the context recognition service does not have to monitor user contexts continuously since users do not naturally change activities and context every second. Therefore, we implemented another energy optimization. It activates the service at every 30 and 60 seconds. Without the screen, and with all sensors, the energy consumption percentage dropped from 40.66% to 20.84%. In addition, with the periodic execution of the service, we were able to reduce the energy consumption by almost 12% for 30s and about 13% for 60s. We also optimized by disabling some sensors such as Wi-Fi and GPS to know which combination of sensors consumes less power. We observed that leaving the GPS in active mode results in a power consumption surge.

6 Discussion

We have evaluated the traditional algorithms to allow their comparisons with the context-based hybrid algorithm. Generally, the context-based hybrid algorithm performed better than the traditional algorithms [Figure 9]. The traditional algorithms show situations where the recommendation without contexts would not give better performance and vice versa. The reason is perhaps that in the process of filtering, user preferences are not filtered according to the contexts where the users have expressed such preferences. Therefore, content that are irrelevant have been included in the recommendations. For example, precisions in Figure 9 e.g. n@5 (a) for target user groups1& 3 are as low as 0.2. The hybrid recommendation approach that combines contextual content and collaborative algorithms produced the overall best results, the same precision in those algorithms increased to 0.7, whilst achieving highest precision of 0.79 @n5 for groups (4&5) and 0.2, 0.3, 0.5 and 0.55 respectively for CBF and CF. All that is required is to identify the user's current contexts and his initial profile to generate recommendations. Additionally, the user satisfaction obtained from users shows that the context-aware hybrid recommendation is feasible in practice.

In terms of power consumption, the context recognition service performed relatively well, though there is a need for improvement through additional optimization.

7 Conclusion

Accessing enormous media items online via smartphones is a significantly different scenario from using desktops to access the same media items. In this article, we have presented CAMR, a framework that recognizes in smartphone user's contexts and physical activities based on smartphone embedded sensors. The framework first monitors and processes the low-level sensor data to derive

high-level contexts or user's physical activities. These contexts and activities are then analyzed to infer situational context, which is at a higher semantic level than the low-level contexts. The framework integrates a contextual user profile service, which relates user's present and past contexts or activities with his preferences to filter online-based media contents for recommendations. Besides, the framework has the capability to track user's interactions with their smartphones, to update his contextual preferences to improve its subsequent recommendations.

Additionally, the evaluation of the recommendation quality of the framework proved that using recognized context and activity information could satisfactorily provide relevant media contents to smartphone users. A user study conducted for 20 users shows that the framework is effective and helpful in discovering interesting online media items based on the user's context and activity information.

Finally, since power consumption by application on smartphone is a very crucial issue, we have evaluated the energy consumption of the framework to understand how we can optimize its energy utilization. This evaluation shows that energy can be significantly conserved if the framework context and the activity recognition service runs at intermittent intervals such as every 30s or 60 seconds. Intermittent switching between GPS or Wi-Fi can also help to conserve power consumption, such as when the user is indoors and with available WiFi connection. In the future, we plan to implement more power consumption optimization techniques to extend the device's battery lifetime. In addition, we would like to evaluate context-aware content based and context aware collaborative filtering recommendations, and compare their performances with that of context-aware hybrid recommendation.

Acknowledgments. The work presented in this paper was partially supported by: Portuguese Foundation for Science and Technology within project FCT/UTA-Est/MAI/0010/2009; the North Portugal Regional Operational Program (ON.2 – O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF).

References

1. Adomavicius, G., Tuzhilin, A.: Towards the next Generation of Recommender Systems: A survey of the State-of-the-art and Possible Extensions. IEEE Transactions on Knowledge and Data Engineering 17(6), 734–749 (2005)
2. Adomavicius, G., Mobasher, B., Ricci, F., Tuzhilin, A.: Context-Aware Recommender Systems. AI Magazine 32(3), 67–80 (2011)
3. De Pessemier, T., Deryckere, T., Martens, L.: Context-Aware Recommendations for User-Generated Content on a Social Network Site. In: Proceedings of the EuroITV 2009 Conference, New York, USA, pp. 133–136 (2009)

4. Meehan, K., Lunney, T., Curran, K., McCaughey, A.: Context-aware intelligent recommendation system for tourism. In: 2013 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops), March 18-22, pp. 328–331 (2013)
5. Chen, A.: Context-Aware Collaborative Filtering System: Predicting the User's Preference in the Ubiquitous Computing Environment. In: International Workshop on Location- and Context-Awareness, Oberpfaffenhofen, Germany, pp. 244–253 (2005)
6. Otebolaku, A.M., Andrade, M.T.: Context Representation for Context-Aware Mobile Multimedia Recommendation. In: Proceedings of the 15th IASTED International Conference on Internet and Multimedia Systems and Applications, Washington, USA (2011)
7. Dong, L., Xiang-wu, M., Jun, L.C.: A Framework for Context-Aware Service Recommendation. In: 10th International Conference on Advanced Communication Technology, ICACT 2008, February 17-20, vol. 3, pp. 2131–2134 (2008)
8. Wenping, Z., Lau, R., Xiaohui, T.: Mining Contextual Knowledge for Context-Aware Recommender Systems. In: 2012 IEEE Ninth International Conference on e-Business Engineering (ICEBE), September 9-11, pp. 356–360 (2012)
9. Otebolaku, A.M., Andrade, M.T.: Recognizing High-Level Contexts from Smartphone Built-In Sensors for Mobile Media Content Recommendation. In: 2013 IEEE 14th International Conference on Mobile Data Management (MDM), June 3-6, vol. 2, pp. 142–147 (2013)
10. Yu, Z., Zhou, X., Zhang, D., Chin, C.Y., Wang, X., Men, J.: Supporting Context-Aware Media Recommendations for Smart Phones. IEEE Pervasive Computing 5(3), 68–75 (2006)
11. Wang, X., Rosenblum, D., Wang, Y.: Context-aware mobile music recommendation for daily activities. In: Proceedings of the 20th ACM International Conference on Multimedia, Nara, Japan, October 29-November 02 (2012)
12. Ostuni, V.C., Gentile, G., Di Noia, T., Mirizzi, R., Romito, D., Di Sciascio, E.: Mobile Movie Recommendation with Linked Data. In: Cuzzocrea, A., Kittl, C., Simos, D.E., Weippl, E., Xu, L. (eds.) CD-ARES 2013. LNCS, vol. 8127, pp. 400–415. Springer, Heidelberg (2013)
13. Kwapisz, J., Weiss, G., Moor, S.: Activity Recognition using Cell Phone Accelerometers. ACM SIGKDD Explorations Newsletter 12(2), 74–82 (2010)
14. Mobasher, B.: Contextual user modeling for Recommendation. In: Keynote at the 2nd Workshop on Context-Aware Recommender Systems (2010)
15. Bobadilla, J., Ortega, F., Hernando, A., Gutirrez, A.: Recommender systems survey. Knowledge-Based Systems 46, 109–132 (2013)
16. Burke, R.: Hybrid Recommender Systems: Survey and Experiments. User Modeling and User-Adapted Interaction, 331–370, doi:10.1023/A: 1021240730564
17. Resnick, P., Neophytos, I., Mitesh, S., Bergstrom, P., Riedl, J.: Grouplens: An open architecture for collaborative filtering of netnews. In: Proceedings of ACM CSCW 1994 Conference on Computer Supported Cooperative Work, Sharing Information and Creating Meaning, pp. 175–186 (1994)
18. Lara, O.D., Labrador, M.A.: A Survey on Human Activity Recognition using Wearable Sensors. IEEE Communications Surveys & Tutorials 15(3), 1192–1209 (Third Quarter 2013), doi:10.1109/SURV.2012.110112.00192.

19. Java EEE, http://www.oracle.com/technetwork/java/javaee/overview/index.html (accessed in October, 2013)
20. Benitez, A.B., Zhong, D., Chang, S.-F., Smith, J.R.: MPEG-7 MDS Content Description Tools and Applications. In: Skarbek, W. (ed.) CAIP 2001. LNCS, vol. 2124, pp. 41–52. Springer, Heidelberg (2001)
21. Zhang, Y., Wang, L.: Some challenges for context-aware recommender systems. In: 2010 5th International Conference on Computer Science and Education Computer Science and Education (ICCSE), pp. 24–27 (August 2010)
22. Java RESTful Web Services http://docs.oracle.com/javaee/6/tutorial/doc/gijqy.html (accessed in October, 2013)
23. PowerTutor, https://play.google.com/store/apps/details?id=edu.umich.PowerTutor (accessed in October 2013)

Measuring the NFC Peer-to-Peer Data Rate

Geoffrey Ottoy, Sam Van Den Berge, Jean-Pierre Goemaere,
and Lieven De Strycker

Abstract. NFC is a relatively new short-range wireless technology. For bidirectional communication between two NFC devices, the NFC Forum specifies the Peer-to-Peer (P2P) standard. Several challenges remain for exchanging data between a mobile Android device and an embedded device. First of all, the current versions of Android implement the NFC P2P specification only partially. A second challenge is the implementation of the NFC P2P protocol stack on the embedded platform. For the developers to be able to make an educated choice on whether to use NFC (P2P) or not, knowledge of the data rate is essential as well. In this article we provide an overview of these challenges. We also create a representative setup with NFC P2P stacks according to the specification on both the mobile and the embedded device. With this setup we are able to measure the NFC P2P data rate and compare it to NFC connection handover to WiFi.

1 Introduction

The motivation of this work has been based on some emerging trends. That is, smartphones will play an increasingly prominent role in people's lives. They will be used to gain access to all kinds of services, both online and at designated terminals such as vending machines, electronic storage lockers or access control terminals. Near-Field Communication (NFC) is one technology that can be used when a smartphone needs to communicate with a such a terminal.

NFC is already applied in several practical cases. For example, in ticketing applications NFC is gradually replacing paper tickets. This is, for instance, the case in Dutch public transportation or in the London Metro, with respectively the

Geoffrey Ottoy · Sam Van Den Berge · Jean-Pierre Goemaere · Lieven De Strycker
DraMCo Research Group
www.dramco.be
e-mail: info@dramco.be
KU Leuven, Campus Gent (KAHO Sint-Lieven),
Faculteit Industriële Ingenieurswetenschappen, Gebroeders De Smetstraat 1, 9000 Gent
e-mail: geoffrey.ottoy@kuleuven.be

L. De Strycker (ed.), *ECUMICT 2014,*
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_9, © Springer International Publishing Switzerland 2014

OV-chipkaart[1] and Oyster Card.[2] A first advantage is that NFC cards can be reused over and over again, which saves on ink, paper and storage [1]. But tickets are also increasingly being stored on NFC-enabled smartphones, which are less prone to loss. In Paris' and London's public transportation, commuters can use their smartphone as a ticket [2, 3]. NFC ticketing systems are furthermore assumed to increase the throughput of the public transportation [1].

Predictions state that in 2016, 13% of US and Western Europe citizens will use their smartphones as a ticket [4]. Manufacturers such as Asus, HTC, Nokia and Samsung have all released NFC-enabled devices. It is expected that 46% of all the smartphones will support NFC by 2016. The only uncertainty is Apple, which has to date not launched any NFC phones.

The NFC Forum[3] regulates the design of all NFC specifications and norms. Started in 2004, it brings together manufacturers and service providers. Among its members are NXP, Nokia, VISA, Samsung, etc. The main goal of the NFC Forum is to ensure interoperability of devices and protocols and thus help in building the so-called *NFC ecosystem*.

NFC Modes of Communication. NFC can be seen as an extension to Radio Frequency Identification (RFID) operation at 13.56 MHz defined by the ISO 14443 standard [5, 6, 7, 8]. This implies that for NFC two types of devices are defined:

- *Passive devices.* These devices –often referred to as *tags*– don't carry a battery and typically have very limited processing power and memory. The functionality ranges of storage of a simple ID (e.g., a classic RFID tag), over storage of (secured) data (e.g., a Mifare card, smart posters), to devices with a cryptographic co-processor (e.g., Java cards). Passive devices draw their power from an RF field generated by an active device.
- *Active devices.* An active device –also called *NFC-enabled* device– has its own power source. This can be either a net supply or a battery. It typically has more processing power than a passive device. Examples of NFC-enabled devices are (some) smartphones and access points.

One of the main differences between RFID and NFC is that with NFC, active devices are also able to communicate with each other. In this case both devices can alternately generate their own RF field when sending, or one of the devices can decide to act as a passive device. Because of the difference in communication between passive and active devices on the one hand and communication between two active devices on the other, the NFC forum has defined three different modes of communication:

- *Reader/writer mode.* This is the "classic" RFID communication. The active device acts as reader to read tags of the ISO/IEC 14443 A/B, Felica, etc.

[1] OV-chipkaart information website: `https://www.ov-chipkaart.nl/`

[2] What is Oyster? Transport for London – Information website: `http://www.tfl.gov.uk/tickets/14836.aspx`

[3] NFC Forum website: `http://www.nfc-forum.org`

- *Peer-to-peer mode (P2P).* In this mode, two NFC-enabled devices can exchange data with one another, e.g., digital business cards, an interesting URL, ... The data rate, however, is small. For this reason, a *connection handover* mechanism is used when large amounts of data have to be transferred. With this mechanism, all necessary parameters to set up a connection over, e.g., WiFi or Bluetooth, are transferred over NFC P2P. After the connection has been established, the NFC communication is terminated.[4]
- *Card emulation mode.* In this alternative to P2P, an NFC-enabled device will act as a traditional RFID tag. The main advantage is that it allows NFC smart phones to easily blend into the market, without the need to change the existing infrastructure (i.e., the existing card readers). However, in this mode the NFC chip of the smartphone communicates directly with a secure element (SE) on the phone (typically a SIM card), which eliminates intervention of the processor and the OS altogether.

The Need for Peer-to-Peer Communication. The number of applications that rely on NFC, and which implement some form of security (e.g., ticketing or mobile payments), is growing. Typically these security measures require bidirectional communication, for instance, advanced authentication protocols such as Idemix [10] or Direct Anonymous Attestation [11]. Currently, when bidirectional communication is required, the phone operates in card emulation mode. This is e.g., the case with *Google Wallet*,[5] or *Isis*.[6] However, this approach has several drawbacks.

Because the communication from the NFC chip is immediately routed to the SE, this element forms a bottleneck. First of all, payment apps are limited to the memory size of the SE. Second, the processing and access times to SEs are higher. A last drawback is that provisioning credentials to a SE is a complex process. In practice, this implies that every *e-wallet app* requires its own SE.

An alternative to using card emulation with SEs is the so-called *software card emulation*. Instead of passing the NFC communication to the SE, it is captured by an NFC service in the OS. This approach breaks the dependency on the SE i.e., credentials can be stored anywhere (e.g., in the phone's user memory, within a trusted execution environment (TEE), even in "the cloud"), but it also allows for several payment applications to use the NFC functionality.

Roland has written a comprehensive article on this subject [12]. His conclusion is that the main reason to go for software card emulation is that it does not require any changes to the existing payment infrastructure (i.e., the vendor terminals and network infrastructure). However, he also states that the most logical choice for payment and ticketing applications over NFC is to go for P2P because it was designed for easy communication between NFC devices.

[4] The connection handover mechanism is standardized in the ISO/IEC 18092 spec. [9]

[5] `http://www.google.com/wallet/`

[6] `https://www.paywithisis.com/`

With regard to software card emulation, BlackBerry is the only company that supports this approach. There have been patches submitted for Android,[7] but none have been merged with the main Android branch. The only way this can be achieved now on Android is by rooting the phone and installing a custom ROM like CyanogenMod;[8] something that would void your phone's warranty.

So why is card emulation used instead of NFC P2P? This is due to the current state of the technology. The NFC P2P specification, defined by the NFC forum, is the standard for bidirectional communication between two NFC devices. Unfortunately, the current version of Android implements this standard only partially.

Contribution. Our work focuses on NFC P2P communication between an Android mobile device and an (embedded) terminal. Android OS is one of the most-used operating systems for smartphones.[9] Moreover, developing apps for Android is easier compared to iOS or Windows 8 because no developer account is required. Also, when needed, root access to the phone can be enabled.

To that end, we will look at the current state-of-the-art, i.e., how is NFC P2P supported in Android OS, what is required to enable NFC P2P? Also on the side of the (embedded) terminal we add support for peer-to-peer communication by using two open-source libraries: `libnfc`[10] and `libnfc-llcp`.[11] Thus we are able to create a practical setup and use it to perform NFC P2P data rate measurements.

Outline. The remainder of this article is structured as follows. We start by giving some insight in the NFC P2P protocol stack (Sect. 2). Here we identify the key components and the current state-of-the art. In the next section (Sect. 3) we highlight the details of our practical setup. Section 4 then discusses the results we obtained. We state our conclusions in Sect. 5.

2 NFC Peer-to-Peer

2.1 Protocol Stack

For two active NFC devices to communicate with each other (e.g., a smartphone and a vending machine), the NFC Forum defines a *peer-to-peer* (P2P) communication mode. NFC P2P requires a stack of several protocol layers as shown in Fig. 1. The

[7] The patch by SimplyTapp adds software card emulation to the Android OS: `https://github.com/CyanogenMod/android_external_libnfc-nxp`

[8] `http://www.cyanogenmod.org/`

[9] According to IDC, Android and iOS Combine for 92.3% of all smartphone operating system shipments in the first quarter of 2013: `http://www.idc.com/getdoc.jsp?containerId=prUS24108913`. Android takes a market share of 75%. This share only includes the smartphone market and not, e.g., tablets.

[10] Project website: `http://code.google.com/p/libnfc/`

[11] Project website: `http://code.google.com/p/libllcp/`

SNEP layer and LLCP layer are defined by the NFC Forum itself. The lower layers are manufacturer and hardware dependent. In the following paragraphs, we will provide a short overview of all the layers.

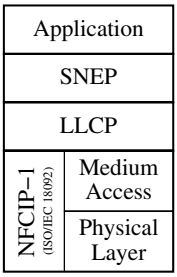


Fig. 1 NFC P2P communication stack

NFCIP-1. The ISO/IEC 18092 specification [9] describes the actual NFC communication i.e., how bits are physically transmitted from one device to another. In [13], the NFC forum has tried to make a more concrete description of the protocol.

The NFCIP-1 protocol works according to an initiator/target model. The initiator sends commands to the target, to which the latter answers. A target can never send data at its own initiative. An initiator is always an active device. Targets can be both active and passive devices. In active communication mode, a target will generate its own RF field when sending data to the initiator. A passive target relies on the RF field generated by the initiator. Through inductive coupling, the target can draw energy from the RF field. The target can be compared to the secondary winding and load of a transformer. By varying the load, the target can send data; the load variations can be detected by the initiator.

LLCP. The Logical Link Control Protocol (LLCP) provides to higher layers a logical connection between two endpoints [14]. This eliminates the inherent asymmetry of the lower NFCIP-1 layers with an initiator controlling the communication and a target device that only sends at a request of the initiator. LLCP creates a so-called *asynchronous balanced mode* (ABM) on top of this mechanism, through which each endpoint has the possibility to start a transmission.

Endpoints connect to the LLCP through *service access points* (SAPs). Typically, each protocol (endpoint) has its own SAP. For NFC, only SNEP has been assigned an SAP (`0x04`).[12]

Implementations on top of LLCP can either use connection-oriented or connectionless communication. Neither of these, however, have a guaranteed delivery time. This makes LLCP unusable for the streaming of audio or video. Other limitations are the fact that there is no possibility to send packets to different SAPs (multicast or

[12] The up-to-date list of assigned SAP addresses can be found at: `http://www.nfc-forum.org/specs/nfc_forum_assigned_numbers_register`

broadcast) and that there is no support for encryption or authentication. This should be implemented by higher layers.

SNEP. The Simple NDEF Exchange Protocol (SNEP) is a request/response protocol for exchanging NDEF[13] messages between a client and a server [15]. To exchange these messages SNEP uses the stable logic connection between two devices, provided by the underlying LLCP layer.

A SNEP message (Fig. 2) consist of four fields:

- The *version* field designates which version of the protocol is being used.
- The *request/response* field indicates the type of the SNEP message. Currently, four request codes are defined. We will only discuss the codes that are relevant for this article.

  `0x00` *CONTINUE*: The client indicates that the server can send another fragment of a fragmented SNEP message.[14] This can only follow a previously sent *GET*. The information field in a *CONTINUE* message is empty.
  `0x01` *GET*: This code is used to request an NDEF message (see later) from the server. The type of NDEF message is determined by the NDEF message that is being sent in the information field of the *GET* request.
  `0x02` *PUT*: Used, when the client wants to sent an NDEF message. The actual message is sent in the information field.

  The possible response codes are defined between `0x80` and `0xFF`. Again we only discuss the codes that are relevant for this article:

  `0x80` *CONTINUE*: Analogous to the request.
  `0x81` *SUCCESS*: Response to a *GET* or *PUT* request. In case of *GET*, the information field contains the requested NDEF message. Otherwise, the information field is empty.

  For a more detailed description and on overview of the other codes, we refer the reader to [15].
- The *length* field (4 bytes) indicates the length of the information field (number of bytes).
- The *information* field is used to send the actual NDEF messages.

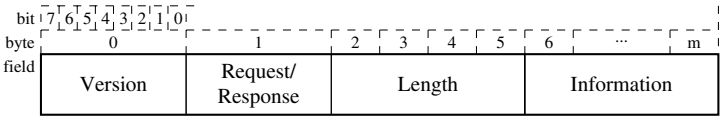


Fig. 2 SNEP message format

[13] NDEF stands for NFC Data Exchange Format. This is discussed later on in this article (p. 115).

[14] This is the case when the information that is being sent, is too large to fit in the information field of a single message.

NDEF Message Format. The NFC Data Exchange Format (NDEF) defines the format of data packets for NFC communication [16]. For P2P communication, when one application sends some information to another, it must encapsulate this information in an NDEF message. An NDEF message consists of one or multiple NDEF records. This can be either a normal record or a short record. Both have a comparable structure with the same fields. For detailed information on the NDEF record format and significance of the data fields, we refer the reader to [16]. Note that the main difference between a short and a normal NDEF record only lies with the allowed payload size.

2.2 Android NFC P2P Stack – API 16 and Higher

The mechanism that handles NFC P2P communication in Android is called *Android Beam* (introduced with API 14). The API provides support for the exchange of NDEF messages, with support for different types of data. There are methods to create the NDEF records and combine them into a message.

To send an NDEF message, the API provides two different approaches. The first one will register a static message to be sent when another NFC-enabled device establishes communication. The second method allows to dynamically create an NDEF message at run time. This is interesting when context-sensitive information needs to be sent.

To receive an NDEF message, an application needs to register the correct intent with the OS; in this case the `NDEF_DISCOVERED` intent. This is important because the *Tag Dispatch System* (Fig. 3) will try to immediately deliver all data received over NFC to the correct application. By using filters, the application can further specify if it is only interested in e.g., plain text NDEF messages. To prevent that a new instance of an application is started, every time an NDEF message is received, an active application can use the *Foreground Dispatch* to process the incoming intents.

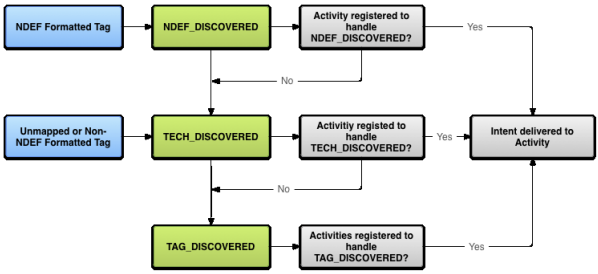


Fig. 3 Android tag dispatch system – Taken from `http://developer.android.com/guide/topics/connectivity/nfc/nfc.html`

If we look at the internal operation of the Android NFC service, one important flaw comes out, more specifically with respect to the NFC P2P specification. The procedure to send an NDEF message (whether it is a static message or composed at run time) is only called when the smart phone receives an *LLCP Link Activation* message. This message is part of the LLCP layer and is *only* sent when two devices are brought close enough to set up a connection. This means that, to be able to send multiple messages, the two devices have to be alternately brought together and moved away again [17]. It is obvious that this is highly impractical. It also implies that currently there is no bidirectional communication possible (with more than one pass), as is required in security protocols.

A possible explanation is that before the release of the NFC P2P specification (by the NFC Forum), Google had its own protocol for exchanging NDEF messages, the NDEF Push Protocol (NPP) [18]. In NPP it is specified that a client has to sever the LLCP connection after sending a NDEF message. It is likely that SNEP was implemented without making any changes to the underlying framework, which explains why this feature is still present.

We note again that this is the case for all Android versions to date.

3 Test Setup

Embedded Platform. With the help of an FPGA development board,[15] we implement a typical embedded platform. This consists of an embedded processor, in this case a 32-bit MicroBlaze™ [21], program memory, an Ethernet interface and an RS-232 interface. The MicroBlaze processor runs embedded Linux, which enables easy application development. Furthermore, it allows us to use open source libraries for NFC communication, as well as standard libraries for data manipulation, networking and input/output in general.

For the RF interface we use the NXP PN532C106 demo board [22]. The heart of this board is formed by the PN532 NFC chip. An antenna and matching network are also provided as well as a power supply circuit. The MicroBlaze processor controls the PN532 chip using the high-speed UART interface. The PN532 and successors such as the PN544 are among the most-used NFC chips in smartphones. The PN532 supports four different modes of operation:

- support for ISO/IEC 14443A (Mifare) and FeliCa as reader
- card emulation as ISO/IEC 14443A and FeliCa tag
- support for ISO/IEC 14443B only as reader
- NFCIP-1, required for NFC P2P; this is the mode that we use

Several open-source libraries exist that offer some functionality required for NFC P2P. One of the main requirements is that the libraries are able to run on an embedded platform. In particular we selected two open source libraries that work with

[15] We use the Xilinx ML605 development board [19], which houses a Virtex 6 FPGA [20].

Linux and that support the PN532. For the NFCIP-1 layer (mainly controlling the PN532) we use `libnfc`.[16] For the LLCP layer we opted for `libnfc-llcp`.[17] On top of that, we have written our own SNEP and NDEF implementation. Currently we only support short NDEF records. With this NFC P2P protocol stack, we are able to exchange messages with an Android smartphone.[18]

Android Phone. As pointed out before, bidirectional P2P communication with Android is currently not possible. To that end, we use a custom NFC P2P stack on the Android phone as presented in [17]. This solution uses the Java Native Interface (JNI) to create a link from Java applications to code written in e.g., C or C++. This code is typically time-critical code, or low-level driver code. Our replacement NFC service implements the NFC P2P stack as defined by the NFC Forum. This includes the NFCIP-1 and LLCP (both written in C) coupled through JNI to a Java API that is usable by applications that require NFC P2P.

Note the with this approach, the phone needs to be rooted because the access to the NFC chip is restricted to the root user or software that is signed by Google.

Another remark is that at the moment, the only solution that does not require root access or OS modification, is connection handover (standardized in ISO/IEC 18092 [9]). The drawback here is that this requires an extra wireless interface and that there is an increased connection setup time. The latter makes that this mechanism is only interesting when a certain amount of data will be exchanged.

Complete Setup. The embedded platform is connected via the on-board Ethernet connection to a WiFi router. With this setup we are able to evaluate both the data rate of NFC P2P, and the communication speed of NFC connection handover to WiFi (TCP/IP).

The complete setup is shown in Fig. 4. In the case of connection handover, the embedded platform transmits an NDEF message to the mobile device. This message contains the handover information that is required to set up the connection over WiFi, e.g., network ID, the platform's IP address, the TCP port to connect to.

As SNEP is a request/response protocol: a new message will only be sent when a *SUCCESS* response on a previous message has been received. We will use this mechanism to measure the time needed to send a certain number of bytes from the platform to the smartphone. At the moment when our embedded platform sends a SNEP *PUT* request, a time stamp is taken. The SNEP *PUT* message contains an NDEF record which in turn contains the actual payload. When the smartphone receives this message successfully, it will reply with a SNEP *SUCCESS* response. Upon receipt of this message at the embedded platform, a second time stamp is taken to measure the time needed to send a message with a certain payload size. Because the NFC stack on the embedded platform is influenced by the load on the OS, communication times will vary (several ms). To obtain an averaged result, we performed

[16] Project website: `http://code.google.com/p/libnfc/`

[17] Project website: `http://code.google.com/p/libllcp/`

[18] Using the standard "Android Beam" service as well as using our own NFC P2P service.

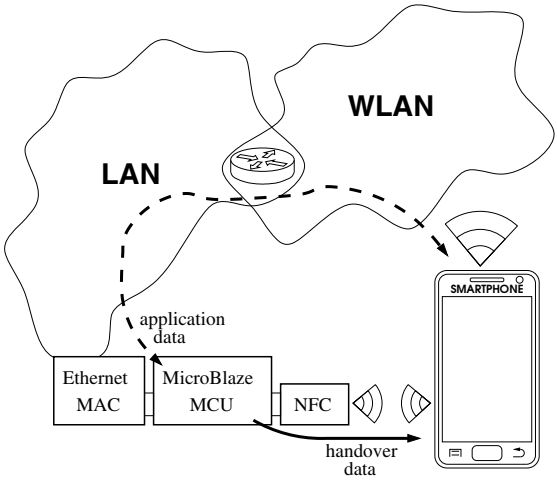


Fig. 4 Test setup

10 measurements for every payload size.[19] To eliminate the connection setup time in our data rate measurements we exchange data over a connection that has already been established.[20] Another simplification is that only short NDEF records are being used. The main reason is that the underlying hardware has a maximum payload length and thus normal NDEF records would have to be parsed over several LLCP messages anyway.

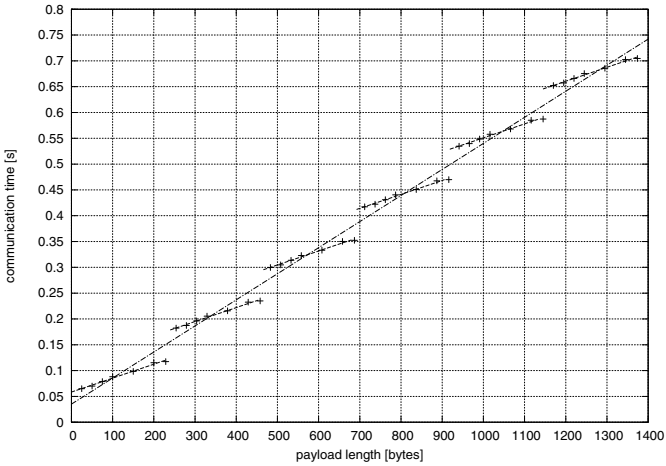
4 Results

NFC P2P Data Rate. We have set out the average time required to send an NFC P2P message (i.e., an NDEF record contained in a SNEP message) for different payload sizes in Fig. 5(a). By dividing the payload size by the time to send the data, we have obtained the effective data rate. This is set out in Fig. 5(b).

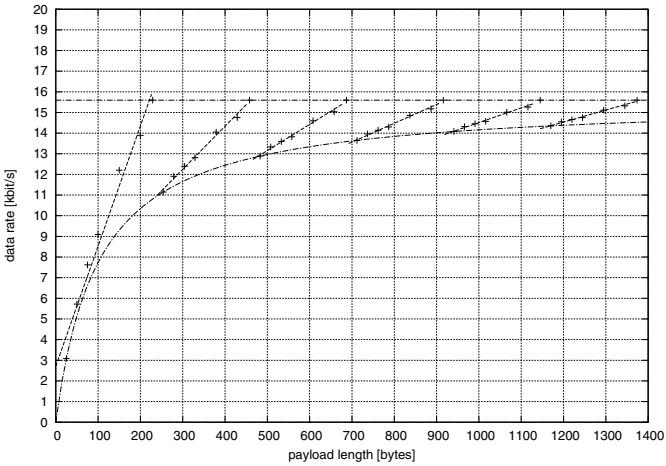
A first remarkable observation is that the time increases linearly in intervals of 229 bytes. Whenever the payload size exceeds a multiple of 229, an extra overhead is introduced. The reason for this is that 229 bytes is the maximum payload size that can be sent with one packet by the underlying layers. This means that when exceeding a payload size of 229 bytes (or a multiple) an extra message needs to be sent, which introduces overhead in the form of headers and the processing of intermediate SNEP *CONTINUE* responses.

[19] We limited ourselves to 10 measurements per payload, because the measurement itself is a tedious procedure that involves resetting the Android application and the embedded NFC stack, as well as "touching" the platform with the smartphone.

[20] Typically connection setup times for NFC are under 100 ms.



(a) Time required to send a number of bytes over an NFC P2P link (in one direction).



(b) Effective data rate for NFC P2P in function of the payload length.

Fig. 5 Speed measurements for NFC P2P communication

This obviously has its influence on the data rate as well. Only for payloads above 1024 bytes, the data rate more or less stabilizes at about 15 kbit/s, with a maximum of 15.6 kbit/s. This is only 3.7% of the physical data rate of 424 kbit/s.

NFC Connection Handover to WiFi. As an alternative to NFC P2P communication, we investigate connection handover to WiFi (TCP/IP). The timing measurements have been performed with the smartphone already connected to the WLAN. This eliminates variations due to the wireless router's response times. The time to

handover the connection has been measured as time difference between sending the handover data and accepting the TCP connection. Over this connection we have send a message of 683 bytes.[21] We have measured this transmission time as well. The connection handover with TCP/IP communication has been repeated 40 times.

We have generated a histogram for the handover timing, the time required for TCP/IP data transmission, and the total communication timing (handover and data transmission). This has been set out in Fig. 6. It is clear that the communication handover is responsible for a large portion of the total communication time. Moreover, it is impossible to do the handover under 0.4 seconds. This is more than the 350 ms for the total communication when NFC P2P is used. The TCP/IP data transmission on the other hand clearly outperforms the NFC P2P communication. This shows that communication handover only pays off when the amount of data to be sent is large enough. More detailed measurements, however, are required to determine the exact tipping point. Specifically for attribute-based credentials it should be investigated if it pays off to use handover, when more attributes are being used.

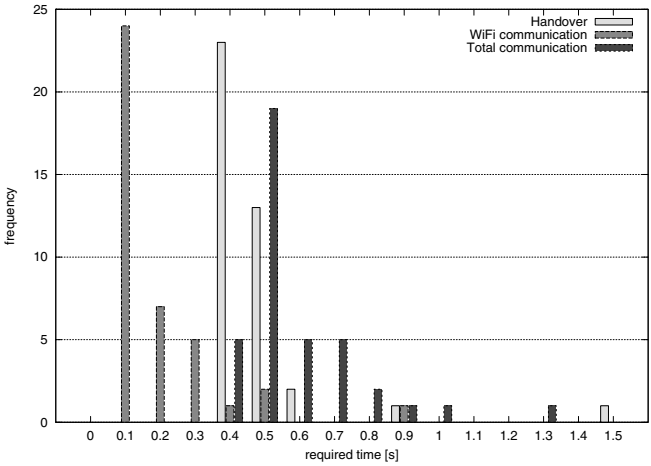


Fig. 6 Communication handover timing

5 Conclusion

As a first practical result, we have been able to determine the effective data rate for NFC P2P communication i.e., 15.6 kbit/s. These are likely the first figures published regarding this subject. However, more measurements are required to determine the

[21] As a practical example we took this value because it is the size of an Identity Mixer [10] credential proof (1536-bit modulus).

effect of the connection set up and the effect of transmitting application data in the other direction as well. Furthermore, when an new Android release implements NFC P2P (according to the specification), it would be interesting to repeat the measurements using this Android version and compare the results with our measurements.

A second conclusion is that NFC P2P is a valid candidate as communication standard when only small amounts of data need to be transmitted (e.g., up to 1024 bytes). Newer versions of the standard (at 848 kbit/s and higher) will result in even shorter communication times. For protocols that require more data transmission, communication handover could be used. However, more detailed measurements are required to determine the exact conditions for which it is beneficial to use either of the approaches. As these are preliminary results, also more research is needed to determine the influence of the protocol overhead, the stack implementation and the OS overhead on the data rate.

We must note that we have only measured the time of setting up a TCP connection and not the time required to connect to a WiFi network. This should deserve some further attention. Other future work could be targeted to comparing connection handover to other technologies such as Bluetooth or GPRS.

One could say that for the further evolution of NFC applications that require P2P communication, the ball is now in Google's court. On the other hand, we assume that eventually they will support the P2P standard. This is a reasonable assumption because 1) an NFC P2P standard is available and 2) applications, other than payment apps, could benefit from this as well.

References

1. NFC Forum, NFC in public transport – white paper (2011), http://www.nfc-forum.org/resources/white_papers/NFC_in_Public_Transport.pdf
2. Clark, S.: Paris commuters to get travel passes that work with NFC phones (2012), http://www.nfcworld.com/2012/01/30/312832/paris-commuters-to-get-travel-passes-that-work-with-nfc-phones/
3. McLean, H.: Transport for London to accept NFC payments from 2012 (2011), http://www.nfcworld.com/2011/07/12/38537/transport-for-london-to-accept-nfc-payments-from-2012/
4. The Point of Sale News, Where Is NFC Going? New Reports Forecast Growth (2012), http://pointofsale.com/20120319953/Mobile-POS-News/where-is-nfc-going-new-reports-forecast-growth.html (date consulted October 10, 2013)
5. ISO/IEC 14443-1:2008 Identification cards – Contactless integrated circuit cards – Proximity cards – Part 1: Physical characteristics
6. ISO/IEC 14443-2:2010 Identification cards – Contactless integrated circuit cards – Proximity cards – Part 2: Radio frequency power and signal interface
7. ISO/IEC 14443-3:2011 Identification cards – Contactless integrated circuit cards – Proximity cards – Part 3: Initialization and anticollision
8. ISO/IEC 14443-4:2008 Identification cards – Contactless integrated circuit cards – Proximity cards – Part 4: Transmission protocol

9. ISO/IEC 18092:2013 Information technology – Telecommunications and information exchange between systems – Near Field Communication – Interface and Protocol (NFCIP-1)
10. IBM Research – Zurich. Specification of the Identity Mixer Cryptographic Library – Version 2.3.2 (2010)
11. Brickell, E.F., Camenisch, J., Chen, L.: Direct Anonymous Attestation. In: Proceedings of the Eleventh ACM Conference on Computer and Communications Security, CCS 2004, pp. 132–145. ACM (2004)
12. Roland, M.: Software Card Emulation in NFC-Enabled Mobile Phones: Great Advantage or Security Nightmare. In: Fourth International Workshop on Security and Privacy in Spontaneous Interaction and Mobile Phone Use (IWSSI/SPMU 2012), p. 6 (2012)
13. NFC Forum. NFC Digital Protocol Technical Specification (2010)
14. NFC Forum. Logical Link Control Protocol Technical Specification (2011)
15. NFC Forum. Simple NDEF Exchange Protocol Technical Specification (2011)
16. NFC Forum. NFC Data Exchange Format (NDEF) Technical Specification (2006)
17. Berge, S.V.D.: Onderzoek en realisatie van P2P communicatie tussen een Android smartphone en een embedded NFC terminal. Master's thesis, KAHO Sint-Lieven – Dept. I.I., Electronics-ICT, Jean-Pierre Goemaere, promotor (2013)
18. Google. Android NDEF push protocol specification (2011), http://source.android.com/compatibility/ndef-push-protocol.pdf
19. Xilinx®. ML605 Hardware User Guide UG534 (v1.8) (2012), http://www.xilinx.com/support/documentation/boards_and_kits/ug534.pdf
20. Xilinx®. Virtex-6 Family Overview DS150 (v2.4) – Product Specification (2012), http://www.xilinx.com/support/documentation/data_sheets/ds150.pdf
21. Xilinx®. MicroBlaze Processor Reference Guide – Embedded Development Kit EDK 14.1 UG081 (v14.1) (2012), http://www.xilinx.com/support/documentation/sw_manuals/xilinx14_1/mb_ref_guide.pdf
22. NXP. AN10609_3 PN532 C106 application note Rev. 1.2 (2010)

A Physical Model for Predicting throughput of Wireless LANs

Mostafa Pakparvar, David Plets, Luc Martens, and Wout Joseph

Abstract. This paper presents an experimentally verified heuristic physical model to predict the throughput of a Wireless LAN link in presence of homogeneous interference of neighboring networks. The model predicts achievable throughput based upon two interference characteristics: transmission rate of the interface and the channel occupancy degree of the interference which is a measure of user activity defined in the paper.

1 Introduction

Large scale growth of wireless networks and spectrum scarcity are introducing more interference than ever. Intensive interference degrades wireless links and may jeopardize seamless connectivity and hence Quality of Service (QoS) offered to the users.

Cognitive Radios (CRs) are becoming a tempting solution to tackle this type of spectrum over-utilization by introducing opportunistic usage of frequency bands that are not heavily occupied by licensed users [1]. The equal regulatory status of wireless terminals on the Industrial Scientific Medical (ISM) band leaves no consideration of a primary user for wireless users. Therefore, interoperability of ISM band networks is becoming a key issue that must be solved.

Wireless networks are designed to tackle homogeneous intra/inter network interference by means of various medium access techniques. The efficiency of interference mitigation techniques of the present technologies has been profoundly assessed in literature. For instance in [2], a large number of experiments prove that

Mostafa Pakparvar · David Plets · Luc Martens · Wout Joseph
Department of Information Technology (INTEC), Ghent University, iMinds
Gaston Crommenlaan 8, Box 201, 9050 Gent, Belgium
e-mail: Mostafa.Pakparvar@intec.ugent.be

L. De Strycker (ed.), *ECUMICT 2014,*
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_10, © Springer International Publishing Switzerland 2014

the throughput of IEEE 802.11 WLAN is highly dependent on the traffic characteristics of other present Wi-Fi networks.

Artificial neural networks (ANN) have also been used for radio parameter adaptation in CR [3, 4]. The ANN determines radio parameters for given channel states with three optimization goals, including meeting the bit error rate (BER), maximizing the throughput and minimizing the transmit power. In [5], it is proposed to use the ANN to characterize the real-time achievable communication performance in CR. Since the characterization is based on runtime measurements, it provides a certain learning capability that can be exploited by the cognitive engine. The simulation results demonstrate good modeling accuracy and flexibility in various applications and scenarios.

Metaheuristics [6] are used for combinatorial optimization in which an optimal solution is sought over a discrete search-space. Evolutionary Algorithms/Genetic Algorithms (GA)s, Simulated Annealing (SA), Tabu Search (TS) and Ant Colony Optimization (ACO) are examples of metaheuristic algorithms. Among the various metaheuristic algorithms, the GA has been widely adopted to solve multiobjective optimization problems and to dynamically configure the CR in response to the changing wireless environment. For instance in [7], Park et al. validate the applicability of GA-based radio parameter adaptation for the CDMA2000 forward link in a realistic scenario with Rician fading. In [8], a GA-based cell-by-cell dynamic spectrum-allocation scheme is investigated achieving better spectral efficiency than the fixed spectrum-allocation scheme. A new solution-encoding technique is proposed to reduce the GA convergence time. In [9], a software testbed for CR with the spectrum-sensing capability is implemented and a GA-based CE to optimize radio parameters for dynamic spectrum access (DSA).

Game theory techniques have also been widely used in the context of cognitive radios. In [10], population game theory has been applied to model the spectrum access problem and develop distributed spectrum access policies based on imitation, a behavior rule widely applied in human societies consisting of imitating successful behaviors. In [11], the authors study the spectrum access problem in cognitive radio networks from a game-theoretical perspective. The problem is modeled as a non-cooperative spectrum access game where secondary users simultaneously access multiple spectrum bands left available by primary users, optimizing their objective function which takes into account the congestion level observed on the available spectrum bands.

Apart from all aforementioned sophisticated methods, there are also methods in literature that derive a physical model for the target parameters based on monitoring physical layer parameters like Received Signal Strength Indicator (RSSI) or Carrier to Interference and Noise Ratio (CINR). For instance, in [12], the authors derive such a physical relation for modeling the throughput of WiMax networks based on the CINR value at the receiver.

In this paper we develop a novel experimentally verified heuristic model to predict and optimize the throughput of wireless terminals in a cognitive network. The model is to be used as the kernel of a cognitive decision engine. The prediction is based on the physical spectrum information. Information is collected passively from

the environment i.e., it does not require any interaction with external/third party networks and resources.

The model is derived from a set of exploratory measurements and is validated with various practical measurements in a pseudo shielded experimental testbed building, w-iLab.t [13]. This paper is organized as follows: in Section 2, the exploratory measurements are described. Section 3, presents the heuristic physical model. Section 4 presents the conclusion of this paper.

2 Exploratory Measurements

After describing the network configuration and the experiment setup, we will elaborate the exploratory measurements we performed for deriving the physical model of the throughput.

2.1 Experiment Description

2.1.1 Network Configuration

All experiments are conducted in a pseudo-shielded testbed environment [13] in Ghent, Belgium. The nodes in the testbed are mounted in an open room (66 m x 20.5 m) in a grid configuration with an x-separation of 6 m and a y-separation of 3.6 m. Figure 1 shows the ground plan of the living lab with an indication of the location of the nodes. Each node has two Wi-Fi interfaces (Sparklan WPEA-110N/E/11n mini PCIe 2T2R chipset: AR9280) and on each Wi-Fi card, two antennas are connected (2x2 MIMO (Multiple-Input Multiple-Output) is supported). Furthermore, a ZigBee sensor node and a USB 2.0 Bluetooth interface (Micro CI2v3.0 EDR) are incorporated into each node.

2.1.2 Experiment Setup

In each of the interference measurements, a set of experiments is executed. Each experiment consists of two phases: a first phase where the QoS of the link under test is recorded without interference, and a second phase where interference is generated in the environment. Phase 2 of each experiment proceeds as follows:(i) Start the interference transmission (ii) Wait 1 second to insure interference is on the air (iii) Start the link under test transmission for n seconds (iv) Stop all data transmissions. The value of n is set to 10 by default unless otherwise mentioned. During the execution of an experiment, the Iperf output stream containing the periodic throughput reports of the receiver is parsed and stored in a database at the experiment controller server of the testbed. For each record, we also log the time stamp of the throughput

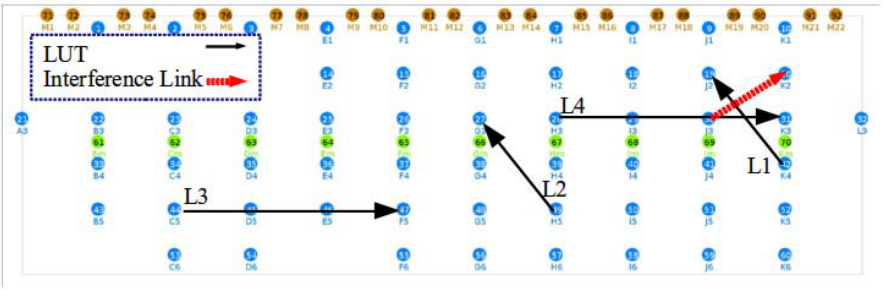


Fig. 1 W-iLab.t living lab test environment (66m x 20.5m) with indication of the nodes and the connection links used for the experiments. Links L1 and L2 are 9.37m long and L3 and L4 are 18m long.

report. Moreover, all physical-layer changes such as channel switching and transmit power modifications together with their time stamps are logged into the database. After running the experiments, the time stamps of the physical- layer changes are used in order to calculate the achieved throughput during each of the experiments in the total set.

During each interference measurement, two connection links are established, as displayed in Figure 1. The first link, always operating on channel 6, is denoted as the link under test (LUT). By default this link is set at the location L1 (see Figure 1) unless otherwise mentioned. The second link is the interfering link, between node 30 and node 20 (see the solid link in Figure 1). The data generation rate of the transmitters is defined as the UDP (User Datagram Protocol) bandwidth of the Iperf application. All Wireless Local Area Network (WLAN) radio interfaces are set to transmit at a certain bit rate (TxRate) and Iperf [14] was used as the traffic generator. The channel occupancy degree (COD) is here defined as the ratio of the data generation rate and the TxRate. There is a transmit buffer that is filled by data at a rate equal to the data generation rate, while the interface transmits through this buffer at a rate equal to the TxRate. The data generation rate is always smaller than or equal to the TxRate.

$$\text{COD}\,[\%] = \frac{\text{Data Generation Rate [Mbps]}}{\text{TxRate [Mbps]}} \times 100 \tag{1}$$

2.2 *Measurement I: Assessment of Channel Overlap and Interference Level Influence on throughput*

The first measurement studies the influence of channel overlap of the interference link on the throughput of the link under test. This measurement consists of 70 iterations each characterized by the interference channel (5 channels) and by the

generated interference level (14 levels) as observed at node 19 (caused by interfering node 30). The channel of the interfering link was varied from channel 6 (full overlap with link under test) to channel 10 (smallest overlap with link under test). For each channel overlap, fourteen different interference power levels [dBm] were observed at node 19: [-51,-52,-53,-54,-56,-57,-59,-60,-62,-63,-66,-70,-71,-72]. This measurement was repeated for 2, 18, and 48 Mbps TxRates for the case of interference on channel 6 and 7 and for 2, 11, 24, and 48 Mbps TxRates for the case of interference on channel 8, 9, and 10 to see if the TxRate influences the throughput on non-overlapping channels. In all iterations, the transmission rate of the LUT equals that of the interference link and the COD of the interference link is equal to 100% (see Equation 1). The value of n (duration of the assessment) was set to 500 seconds in this measurement.

Figure 2 shows the achieved average throughput as a function of the received interference power level, for interference on different channels. Each region is identified by a bit rate which is the TxRate of the LUT and interference link. Note that here the interference COD is equal to 100%. A higher channel number corresponds with less overlap, i.e., for increasing channel separations, interference decreases (LUT operates on channel 6). The interference level on the x-axis of these Figures is defined as the power received from node 30 at node 19 (see Figure 1).

Measurement results in Figure 2 show that the sensitivity of the throughput to the interference channel overlap depends on the transmission rate of the links. Firstly, the interference power level does not affect the throughput when the interference power level is above a certain TxRate-channel dependent threshold: the throughput is not more impacted when the interference power is higher.

The upper regions of Figure 2 reveal high sensitivity of throughput on interference level and interference transmission rate at higher TxRates. Interference on channels 9 and 10 (channel separation of 3 and 4) affects the throughput only when the interference TxRate is above 2 Mbps. For the 48 Mbps case for instance, there is more than 30% throughput reduction for interference power levels above -70 and -60 dBm on channels 9 and 10 respectively. In all cases, at higher interference TxRates, the interference level is a key parameter that influences the achieved throughput.

Therefore, channel overlap between the interference source and the network under study is an important parameter that affects the throughput of WLANs. The results of these measurements suggest that if the cognitive decision engine has to select an overlapping channel, it should select the channel with the least interference level and has lower interference TxRate.

2.3 Measurement II: Joint Assessment of the Interference TxRate and Interference COD Influence on throughput

Since interference COD and interference TxRate are not completely uncorrelated parameters, deriving a decent 2-dimensional model requires joint assessment of the two parameters influence on throughput. Here L3 serves as LUT and the same

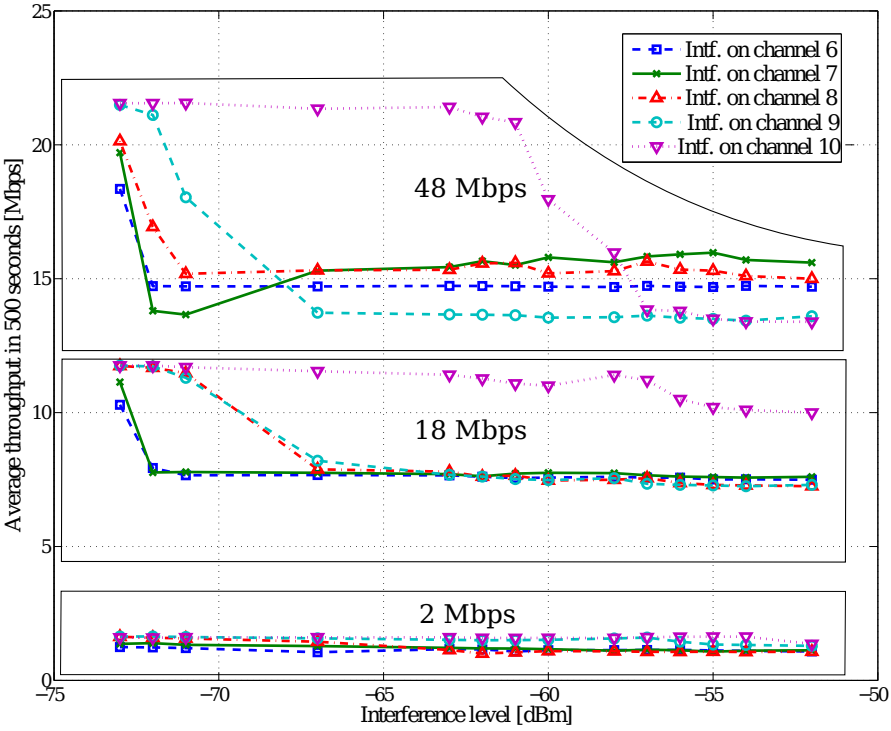


Fig. 2 Average throughput over link under test in a 500 second time frame as a function of received interference power level, for interference with different channel overlaps. Each region has its corresponding TxRate of the LUT and interference link. Intf. COD = 100%. The LUT is on channel 6.

selection of interference link as described in previous measurements is used. We measure the average achievable throughput in a ten seconds interval with interference TxRates of [2 11 18 24 36 48 54] and interference COD values starting from 0 to 100% in steps of 6.25%. All links are set to channel 6 and the transmit power of all terminals is set to 20 dBm. The TxRate of the LUT is set to 54 Mbps and its COD is set to 100% percent such that the maximum achievable throughput is measured.

Sample points in Figure 3 show the results of validation measurement on link L1. Looking at the results of this measurement unveils the dependency of the interference COD and TxRate when predicting the throughput value. This is more investigated in next Section.

3 Development of the Model

3.1 Construction of Models

Data transmission rate of the Wi-Fi interface and the user activity are the key parameters that determine the spectrum utilization by any terminal. For a fixed amount of data transmission, higher transmission rate (TxRate) means less spectrum occupancy in time. Hence, an active user with low transmission rate occupies the spectrum to an extent that users of other networks in the vicinity experience a much lower than expected throughput.

The model predicts the achievable throughput (T) based on the interference characteristics. Parameters that characterize the interference are: COD of the interference, interference TxRate, interference frequency channel. All parameters are obtainable using the monitor mode of IEEE 802.11 interfaces [15] and packet tracing applications like libtrace [16] and tcpdump [17]. Extension to all Wi-Fi channels is feasible by periodic monitoring of all channels or by exploiting more WLAN interfaces at different locations of the network each operating on a single channel.

Due to the dependency of the interference parameters, deriving a decent 2-dimensional model to predict the throughput is crucial. The observation of the throughput behavior in sample points of Figure 3 shows that depending on the interference TxRate, the achieved throughput increases exponentially after a certain interference COD threshold. Therefore, by applying unit step function to the exponential models and partitioning the space into linear and exponential regions, we come up with an accurate model. This is evidenced by the results of our suggested model in Equation 2 below:

$$T(COD_{Intf.}, TxRate_{Intf.}) = (a_0 e^{-b.COD_{Intf.}})u(90 - COD_{Intf.} - r.TxRate_{Intf.}) + (a_0 e^{-b(90 - r.TxRate_{Intf.})})u(COD_{Intf.} + r.TxRate_{Intf.} - 90) \quad (2)$$

where u(t) is the unit step function, $TxRate_{Intf.}$ denotes interference TxRate, $COD_{Intf.}$ denotes interference COD, and the coefficient parameters should be optimized for the intended link and the intended channel. The nonlinear least square optimization results for measurements on link L3 assigns an a_0 value of 23.23, $b = 0.02$ and $r = 0.5$. These values are obtained by exploiting all the sample points of the measurement II. Figure 3 illustrates this realization of the model for sample points of link L3. Note that the configuration of the step function arguments are optimized manually by looking up the results of measurement II where threshold of the aforementioned throughput exponential behavior increases linearly from a COD of 90% for TxRate of 2 Mbps up to the COD value of 60% for TxRate of 54 Mbps. The obtained R^2 (coefficient of determination) value is 0.9425 and the Root Mean Square Error (RMSE) value is 1.34 which show the accuracy of the model.

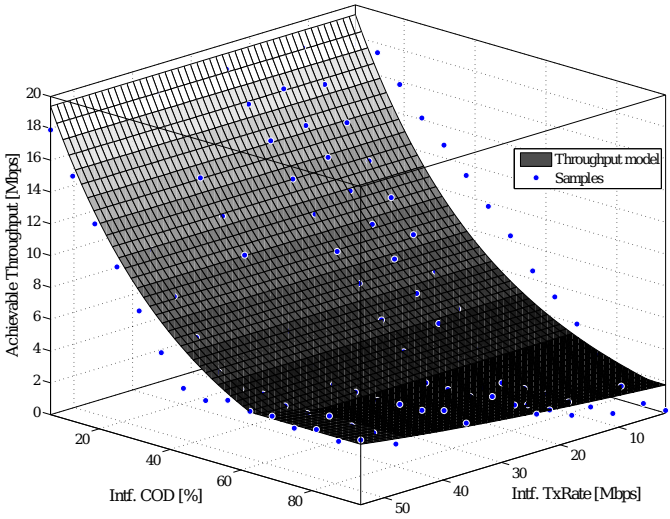


Fig. 3 The 3D demonstration of the throughput model realized at L3

3.2 Validation Measurements

The objective of these measurements is to verify the 2-dimensional model of Equation 2 for different sender-receiver distances and different link locations. To this end, different combinations of LUT terminals with various distances of sender-receiver were used while the interference link was fixed. These links are illustrated in Figure 1: link L2 is a replication of L1 while L4 is a replication of L3 at another location. Except for the LUT terminal locations, the measurement configuration is identical to measurement II described in Section 2.3.

3.3 Validation Results

The model obtained by using measurements at link L3 is assessed with measurements on links L1, L2 and L4. The results are reported in Table 1. All R^2 values are above 0.89 and the RMSE does not exceed 1.7 Mbps. The maximum deviation is 3.3 Mbps which is acceptable when compared to the achievable throughput range of 20 Mbps. Hence, the model optimized for a single link location can serve to predict the throughput at other links with an acceptable accuracy. This measurement setup was limited to the node locations and propagation characteristics of the testbed. However, for environments with more diverse propagation characteristics and larger sender-receiver distances, a correction parameter can be added to the model to compensate the throughput drop due to more attenuation. The simplicity of this model enables the decision engine to devise it with a small number of

Table 1 Statistics of the verifying measurements at various links with the model obtained at L3

Link	RMSE [Mbps]	Max. Deviation [Mbps]	R^2
L4	1.65	3.31	0.8987
L3	1.34	2.93	0.9425
L2	1.59	2.57	0.9056
L1	1.59	2.66	0.9059

sample points. Figure 4 shows the RMSE and R^2 of the model realized with varying number of sample points. Every realization is a set of measurements with all possible $TxRate_{Intf.}$ values and $COD_{Intf.}$ values uniformly distributed between 0% and 100%. The R^2 value is always more than 0.94 and the RMSE is always less than 1.35 Mbps. These statistics show that the model can be devised accurately with 42 samples.

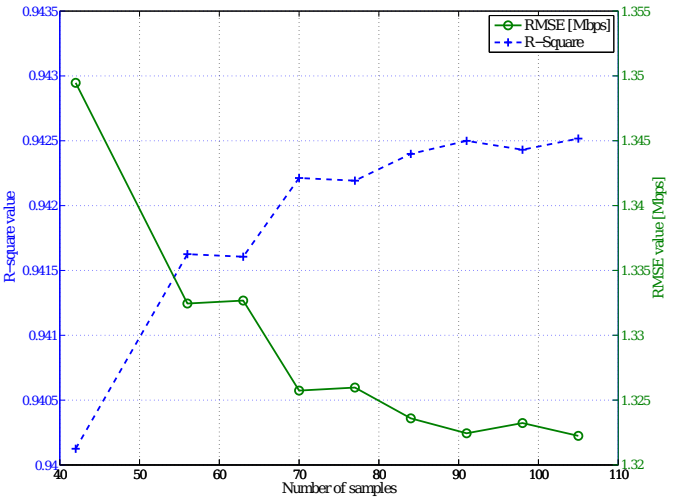


Fig. 4 Statistics of the model for realizations with varying number of sample points

4 Conclusion and Future Work

Based upon numerous exploratory measurements, a novel physical throughput model accounting for interference channel occupancy degree ($COD_{Intf.}$) and interference transmission rate ($TxRate_{Intf.}$) was devised and verified in a pseudo-shielded testlab environment.

Future research on this framework includes comparison of the performance of the cognitive engine based on the proposed model with smarter higher level algorithms, and integration of this type of physical modeling with other decision algorithms to achieve more efficient algorithms. This model was devised in a pseudo-shielded environment, therefore extension to office and industry environments can be the topic of further research.

Acknowledgment. This work was supported by the iMinds-ICON QoCON project, co-funded by iMinds, a research institute founded by the Flemish Government in 2004, and the involved companies and institutions. The research is also partly funded by the Fund for Scientific Research - Flanders (FWO-V, Belgium) project G.0325.11N.

References

1. Mitola, J., Maguire Jr., G.Q.: Cognitive radio: making software radios more personal. IEEE Personal Communications 6, 13–18 (1999)
2. Plets, D., Pakparvar, M., Joseph, W., Martens, L.: Influence of intra-network interference on quality of service in wireless lans. In: Proc. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB 2013) (June 2013)
3. Hasegawa, M., Tran, H.-N., Miyamoto, G., Murata, Y., Kato, S.: Distributed optimization based on neurodynamics for cognitive wireless clouds. In: IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2007, pp. 1–5 (2007)
4. Zhang, Z., Xie, X.: Intelligent cognitive radio: Research on learning and evaluation of CR based on neural network. In: ITI 5th International Conference on Information and Communications Technology, ICICT 2007, pp. 33–37 (2007)
5. Baldo, N., Zorzi, M.: Learning and adaptation in cognitive radios using neural networks. In: 5th IEEE Consumer Communications and Networking Conference, CCNC 2008, pp. 998–1003 (2008)
6. Blum, C., Roli, A.: Metaheuristics in combinatorial optimization: Overview and conceptual comparison. ACM Comput. Surv. 35, 268–308 (2003)
7. Park, S.K., Shin, Y., Lee, W.C.: Goal-pareto based nsga for optimal reconfiguration of cognitive radio systems. In: 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications, CrownCom 2007, pp. 147–153 (2007)
8. Thilakawardana, D., Moessner, K.: A genetic approach to cell-by-cell dynamic spectrum allocation for optimising spectral efficiency in wireless mobile systems. In: 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications, CrownCom 2007, pp. 367–372 (2007)
9. Kim, J.M., Sohn, S.H., Han, N., Zheng, G., Kim, Y.M., Lee, J.K.: Cognitive radio software testbed using dual optimization in genetic algorithm. In: 3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications, CrownCom 2008, pp. 1–6 (2008)
10. Iellamo, S., Chen, L., Coupechoux, M.: Proportional and double imitation rules for spectrum access in cognitive radio networks. Computer Networks 57(8), 1863–1879 (2013)
11. Elias, J., Martignon, F., Capone, A., Altman, E.: Non-cooperative spectrum access in cognitive radio networks: A game theoretical model. Computer Networks 55(17), 3832–3846 (2011)

12. De Bruyne, J., Joseph, W., Verloock, L., Olivier, C., De Ketelaere, W., Martens, L.: Field measurements and performance analysis of an 802.16 system in a suburban environment. IEEE Transactions on Wireless Communications 8(3), 1424–1434 (2009)
13. S.B., et al.: Federating wired and wireless test facilities through emulab and omf: the ilab.t use case. In: Proceedings of TridentCom 2012 (2012)
14. Iperf - the TCP/UDP bandwidth measurement tool, `http://iperf.fr/` (accessed July 19, 2013)
15. IEEE standard for information technology Telecommunications and information exchange between systems local and metropolitan area networks Specific requirements part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. IEEE Std 802.11-2012 (Revision of IEEE Std 802.11-2007), pp. 1–2793 (2012)
16. Libtrace library for trace processing, `http://research.wand.net.nz/software/libtrace.php` (accessed July 19, 2013)
17. tcpdump, a powerful command-line packet analyzer, `http://www.tcpdump.org/` (accessed July 19, 2013)

Study of New Organic Field Transistors for RFID, Optoelectronic and Mobile Applications

Marius Prelipceanu and Adrian Graur

Abstract. We present our results in developing processes and new materials for realization of new organic transistor which are promising for optoelectronics and radio frequency identification (RFID) applications. In this report we discuss the films morphology, profilometry and the field-effect transistor (FET) performances of pyrrolo phenanthroline derivatives (RA). We consider that because of π-conjugation which is extended and good alignment of molecules, the pyrrolo phenanthroline devices exhibited hole mobilities of up 0.031 $cm^2 V^{-1} s^{-1}$. The performance of these devices can be adequate for construction of 135-kHz RFID or high resolution display.

1 Introduction

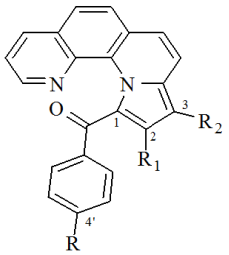
In the last years, studies in organic semiconductors have grown because of their multiple applications in optoelectronic devices and radio frequency identification (RFID) tags [1-9]. These materials have a few advantages compared to conventional inorganic electronics. We can mention here a good and easy processability, a good chemical control concerning the injections of the charge, a good fit in with plastic substrates, and reduced production price [10-13]. It is important to have particular molecular ordered architectures in this material, in order to obtain high carrier mobility. Good alignments of molecules in established orientation is fit in for intermolecular charge migration and also for a productive charge transport, necessary for a new generation of device [14-16]. During the research of organics, a lot of studies were made to obtain highly crystalline films [17-18]. The FETs, RFID and organic emitting diodes characteristics likewise, depend on the film morphology [19]. Here we report on the OFET particularities of pyrrolo compounds [7,8] including correlation between morphology structure and performance of electrical measurement.

Marius Prelipceanu · Adrian Graur
Faculty of Electrical Engineering and Computer Science,
"Ştefan cel Mare" University Suceava, Romania
e-mail: {mprelipceanu,adriang}@eed.usv.ro

L. De Strycker (ed.), *ECUMICT 2014*,
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_11, © Springer International Publishing Switzerland 2014

2 Materials and Methods

Fig.1 shows the structure of pyrrole derivatives. Their synthesis is presented in literature [9-13].



RA1: R= NO_2, R_1=R_2= $COOCH_3$
RA2: R= Cl, R_1=R_2= $COOCH_3$

Fig. 1 Chemical structures of examined compounds

The structures of RA1 and RA2 compounds were studied first by NMR spectroscopy. The signals of the protons have been assigned in accord with the current data presented in literature [9-14]. Therefore, 13C nmr spectra for RA1-RA2 compounds were improved and results certified the chemical structures [15-17].

The electrical properties of the FETs were measured in vacuum. We used a measurement system formed from two Keithley source meters. For surface structure characterizations, the semiconducting materials were spin coated onto quartz-glass substrates simultaneously with the FET devices. The surfaces morphology of the films have been obtained utilizing an AFM, working in non-contact method. The AFM measurements were made in air at ambient temperature.

The thickness and index of refraction of films have been obtained used an Ellipsometer who works at 632,8 nm wavelength. The thickness of the thin films has been compared with data obtained with a Dektak Profilometer, which has the ability to measure the structure down to a few nm. UPS measuring have been effectuated at BESSY (Berlin) [7]. A toroidal grating monochromator was used (range of photon energy - 5 to 190 eV) [7]. The photoelectrons were added together with a special spectrometer system at room condition. Pending the experiment the pressure had a constant value, 2 x 10^{-10} mbar [7].

3 Results and Discussions

We have previously reported UPS measurement of compounds RA1(a) and RA2 (b) and their mathematical simulations for two incident photon energies [7]. We found a good arrangement amongst experimental and theoretical results [7, 15-18]. In the case of our two compounds, the emission is producing especially by π-orbitals [7, 8].

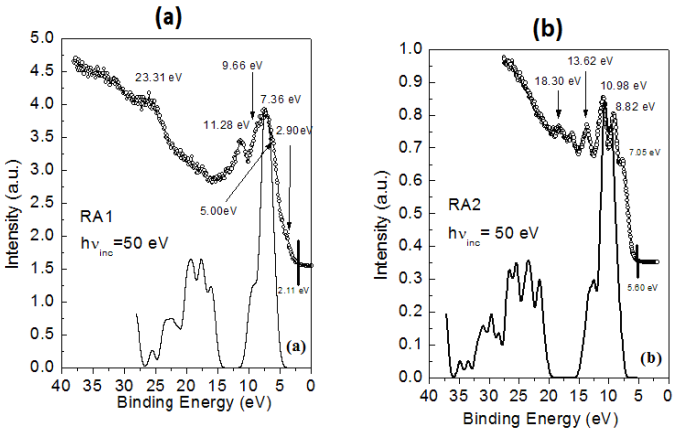


Fig. 2 UPS valence band spectra of compounds RA1 (a) and RA 2 (b) at 50 eV incident photon energy. Experimental-symbol line; simulated spectra-full line and energetic diagram for RA derivatives.

The HOMO levels were established utilizing the onset potential of the oxidation process of RA1 and RA2 compounds [7-8] and this are comparing with the UPS results. The onset of oxidation appears at 0.9 V and 0.8 V, who fulfill to E_{HOMO} values of −5.25 eV and −5.40 eV determined and shown by UPS measurement. Thus the HOMO levels of the active layer fit well with work functions of gold used as electrode. These correspond to a good hole injection in device. FETs were manufactured utilizing the bottom contact geometry (Figure 3c).

The channel's length is $L = 100$ μm and the width is $W = 500$ μm. The top electrodes (source and drain) were fabricated from gold coated by thermal evaporation by means of precise shadow mask. Silicon oxide was used as a dielectric (150 nm).

The thin films were spin coated the organic semiconductor onto a cleaned SiO_2, glass and quartz-glass surfaces and device from a solution 5% dimethylformamide (DMF) in the meantime, in order to study the morphology of layers.

Thus, the pyrrolo phenanthroline derivatives are organic semiconductors, which are enabling for the solution processing. They can dissolve well in organic solvents (DMF, chloroform or chlorobenzene).

We found that such compounds can be used for spin-coating, stamping and printing. In Table 1 are shown the solubilities of phenanthroline derivatives in different organic solvents solutions.

Table 1 Stability and solubility of phenanthroline derivatives in organic solvents

Solvent solution	phenanthroline derivatives
Acetone	Partial soluble/instable in solution/weak films
Chloroform	Soluble/instable in solution/fair films
Ethanol	Partial soluble/ instable in solution/fair films
Chlorobenzene	Soluble/stable in solution/fair films
Dimethylformamide	Soluble/stable in solution/ good films

We fabricated test chips made by doped silicon with an aluminum metallized back serves like the gate electrode, in order to obtain a good electrical characterization of the organic semiconductors (see Fig. 3b). The FET device is completed by spin-coating the organic material. After deposition the test chip was investigated by microscopy methods in order to observe the quality of electrodes and to measure channel length and width. The thickness determinated by AFM and Ellipsometry was in range of 10 - 100 nm. (Figure 3a).

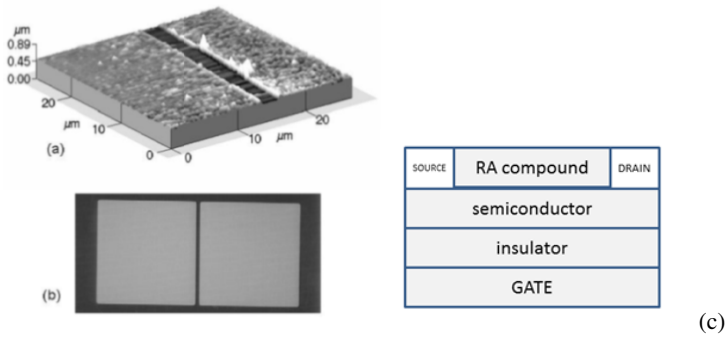


Fig. 3 AFM Electrodes investigation (a) AFM investigations of transport channel, (b) top view of chip. FET bottom contact geometry (c).

The electrical particularities of OFETs depend strongly on the molecular alignment in the film and on the injection of charge from the gold contacts [17-18]. The AFM micrographs of the phenanthroline films described morphologies which vary with spin speed (see Figure 4).

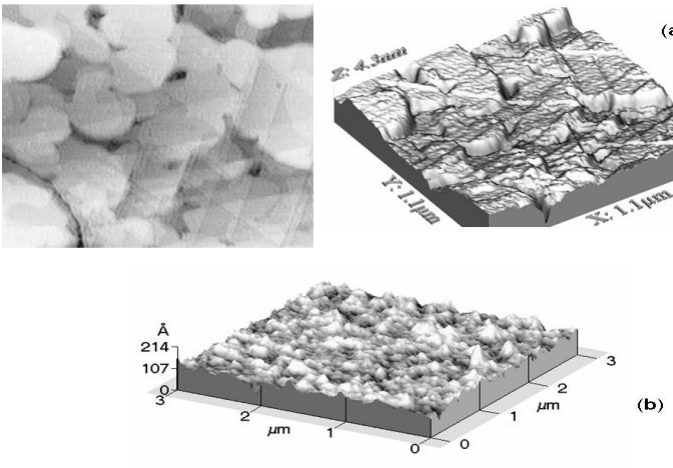


Fig. 4 Surface morphology of phenanthroline derivatives films (a) 3000 rot/min speed spin, (b) 1500 rot/min speed spin

By AFM investigation, was found the preferred crystal orientation in all phenanthroline films. This is due, probably, because of spin-coating direction. The grains are big in dimension, in range of 20 – 160 nm. They are isolated by little grain boundaries.

We observe smooth surface in all investigated films. Root mean square (RMS) roughness determined by ellipsometry and profilometry measurements is in range of 1 nm and 3 nm. At higher speed in spin-coating process, the RMS values which we achieved denote a smoother surface. The typical relief of a phenanthroline derivatives film on a SiO_2 layer is shown in Fig. 4. The layers thick films were investigated by X-ray diffraction. The spectra does not show Bragg peaks because, probably, an insufficient crystallinity of the layers. Also this observation can confirm that domains with different inclination angles are present in the film.

The thin-film transistors of RA1 – RA2 presented typical p-channel characteristics. At a negative bias applique, the drain–source current scaled with the negative gate voltage because of the grown number of charge carriers, holes in our situation. The produced curves at various gate biases and the transfer curves at steady V_D for the phenanthroline derivatives films are shown in Figure 5. The output characteristics show a good saturation region.

We calculate the field-effect mobilities in the saturation regime at a drain voltage of $V_D = -50$ V, the capacitance of the SiO_2 insulator C_i is considered 12.3×10^{-9} F cm^{-2} and the value of V_G is the gate voltage and V_T threshold voltage are from experiment.

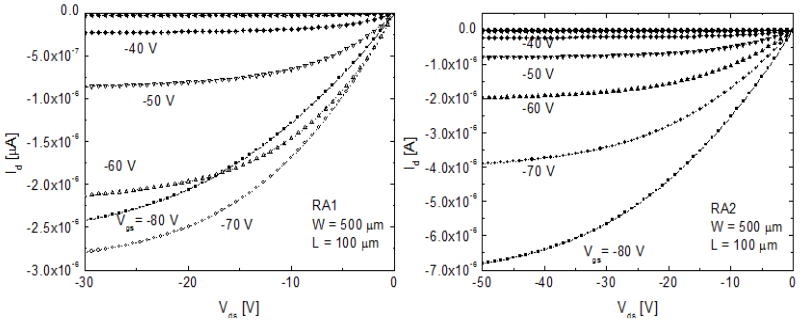


Fig. 5 Output characteristics of the RA1-2 devices at various gate biases

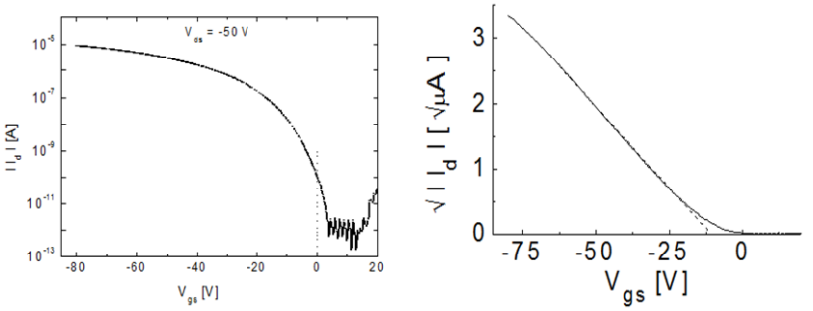


Fig. 6 Plots of the transfer at steady $V_D = -50$ V and semilogarithmic and plot of $(-I_D)^{1/2}$ versus V_G

The FET devices were determined in air and the results are presented in Table 2.

Table 2 Field-effect mobility (μ_{FET}) and threshold voltage (V_{th}) for RA devices at 25^{0}C substrates temperature

Material	T_{sub}[^{0}C]	μ_{FET} [$cm^2V^{-1}s^{-1}$]	I_{on}/I_{off}	V_{th} [V]
RA1	25	0.008-0.006	10^7	-18 to -14
RA2	25	0.031-0.020	10^7	-14 to -11

The phenanthroline derivatives devices showed hole mobilities of in range of 0.0006 – 0.031 cm2 V^{-1} s^{-1} and threshold voltages of −18 V and −11 V. The performances of the RA derivatives devicess may be due to the right fit amongst the ionization potential of each phenanthroline derivatives and the work function of the gold electrodes. To obtain the smaller channel length of device in order to increase the speed function, the gold contacts (source and drain) must cover the gate contact, but in this way we get large capacitance. To understand the impact of

this capacitance on performance, measurements were performed on devices with 300 μm overlap (fig. 7). As expected, the overlap capacitance causes substantial roll-off at 10 kHz, which is reasonable for 135-kHz RFID.

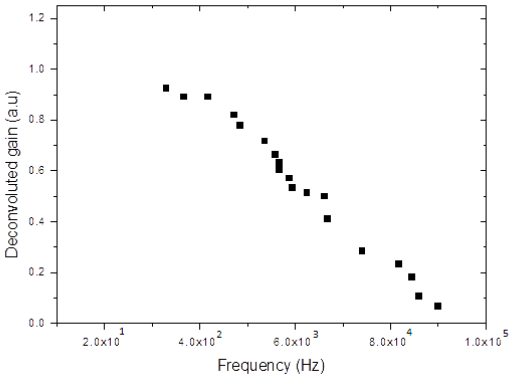


Fig. 7 AC performance of devices with 300 μm overlap

The FET particularities of OFETs depend, in the first case, on the molecular alignment in the film and on the injection of holes from the electrodes. We assume that the molecules are aligned perpendicular to the substrate and the π–π superposition amongst close molecules is increased at maximum and carrier transport can exist. The rigidity and planarity of RA compounds can give a high crystallinity and vertical alignment into the reasonable and efficient molecular orientation. In principle, for such organic compounds that present great mobilities at the highest substrate temperatures because the grain dimension tends to grown and the number of grain boundaries tends to decline with increased temperature [19]. Unfortunately, the OFET devices made from phenanthroline derivatives showed a weak stability after 15 days in air.

4 Conclusions

We have studied new semiconducting organic materials containing different radicals group and manufactured OFETs utilizing those materials as active layer. Because of the expanded π-conjugation, productive charge injection, and right alignment of the molecules, the phenanthroline devices showed hole mobilities in range of 0.0006 - 0.031 $cm^2 V^{-1} s^{-1}$. In phenanthroline films, the carriers, holes in our case, move effectively, and thus the FET particularities are efficient, when the molecular orientation favors π–π superposition and the grains are big in dimensions and isolated by little grain boundaries. The molecules aspired to align approximately perpendicular to the substrate, which favors π–π overlap between close molecules. Furthermore betterment in FET performance may be possible by improving the manufacturing conditions – using various substrates temperatures and the surface treatment in order to improve the orientation of molecules.

Our new organic transistors are promising for construction of integrated circuits on thick flexible polyester substrate, using methods compatible with printing processes for mass production. These circuits can be used in optoelectronics, mobile and radio frequency identification (RFID) applications.

We can mention here few advantages as a good and easy processability, a good chemical control concerning the injections of the charge, a good fit in with plastic substrates, and reduced production price

Acknowledgments. This work was supported by the Project Q-DOC - Improving the quality of doctoral studies in engineering sciences to support the development of the knowledge society – Contract POSDRU/107/1.5/S/78534"co-funded from European Social Fund through Sectorial Operational Program Human Resources 2007–2013. We gratefully acknowledge Prof. Dr. Liviu Leonte and Dr. R. Danac, to provide us the phenanthroline derivatives.

References

1. Burroughes, J.H., Bradley, D.D.C., Brown, A.R., Marks, R.N., Mackey, K., Friend, R.H., Burns, P.L., Holmes, A.B.: Nature 347, 539 (1990)
2. Friend, R.H., Gymer, R.W., Holmes, A.B., Burroughes, J.H., Marks, R.N., Taliani, C., Bradley, D.D.C., Dos Santos, D.A., Brédas, J.L., Lögdlund, M., Salaneck, W.R.: Nature 397, 121 (1999)
3. Shim, H.-K., Jin, J.: Adv. Polym. Sci. 158, 193 (2002)
4. Katz, H.E., Bao, Z., Gilat, S.L.: Acc. Chem. Res. 34, 359 (2001)
5. Horowitz, G.: Adv. Mater 10, 365 (1998)
6. Babel, A., Jenekhe, S.A.: Macromolecules 36, 7759 (2003)
7. Prelipceanu, M., Prelipceanu, O.S., Leontie, L., Danac, R.: Physics Letters A 368(3-4), 331–335 (2007)
8. Prelipceanu, M., Prelipceanu, O.S., Tudose, O.G., Leontie, L., Grimm, B., Schrader, S.: Materials Science in Semiconductor Processing 10, 77–89 (2007)
9. Zugravescu, I., Petrovanu, M.: Rom. Acad. Publ. House, Bucharest (1987) (in Romanian)
10. Druta, I., Andrei, M., Aburel, P.: Tetrahedron 54, 2107 (1998)
11. Druta, I., Dinica, R., Bacu, E., Humelnicu, I.: Tetrahedron 54, 10811 (1998)
12. Dinica, R., Druta, I., Pettinari, C.: Synlett 7, 1013 (2000)
13. Danac, R., Rotaru, A., Drochioiu, G., Druta, I.: Heterocyclic Chem. 40 (2003)
14. Ito, K., Isobe, T., Sone, K.: J. Chem. Phys. 31(861), 283 (1959)
15. Blears, D., Danyluk, S.: Tetrahedron 23, 2927 (1967)
16. Carman, R., Hall, J.: Aust. J. Chem. 17, 1354 (1964)
17. Rosenberger, H., Pettig, M., Madeja, K., Pehk, T., Lippmaa, E.: Org. Magn. Reson. 2, 329 (1970)
18. Meng, H., Zheng, J., Lovinger, A.J., Wang, B.-C., Patten, P.G.V., Bao, Z.: Chem. Mater. 15, 1778 (2003)
19. Ando, S., Nishida, J.-I., Tada, H., Inoue, Y., Tokito, S., Yamashita, Y.: J. Am. Chem. Soc. 127, 5336 (2005)

Applicability of Amdahl's Law in Multisession TCP/IP Communication

Radu-Cezar Tarabuta, Alin Potorac, Doru Balan, and Adrian Graur

Abstract. In this paper the authors are proposing to evaluate whether Amdahl's law is valid in the transmission's data using TCP/IP protocol, based on existing similarities. In our study the parallel processors cores are replaced with parallel communication sessions. The paper is demonstrating the possibility to extend the use of Amdahl's law to multisession communications as a different method for QoS analysis in specified conditions.

1 Introduction

The Amdahl's Law was proposed back in 1967 and it basically defines the possibility to predict a system improvement in terms of processing speed while components or parts of the systems are individually improved. The starting approaches were related with parallel computing techniques and offer a solution to evaluate how much a program execution is speeded up when multiple processing cores are used instead of a single one. Obviously, there is not a linear dependency and a lot of parameters are usually involved.

Amdahl found that the speedup in a distributed process can be computed starting from the time parameters associated with serial and parallel processing over multiple processors with a simple formula which eventually can be adjusted to include the impact of different other factors.

The researches described in this article present a possible analogy between Amdahl's approach and a computer network communications improvement relative to the bandwidth growth based on session's multiplication.

This law basically relates to distributed processing but can be a perfect analogy to what happens in data communication. Basically the Amdahl's law says that if we have a process that takes 20 hours of execution on a single processor core and we can manage this process to "break" in several individual sub-processes that separately would take an hour and we manage to allocate these sub-processes to 20 cores, execution of the whole process would take an hour.

If N is the number of processors, s is the amount of time spent (by a serial processor) on serial parts of a program and p is the amount of time spent (by a serial

Radu-Cezar Tarabuta · Alin Potorac · Doru Balan · Adrian Graur
University "Stefan cel Mare" Suceava, Romania
e-mail: tarabutaradu@yahoo.com, {alin,dorub,adriang}@usv.ro

L. De Strycker (ed.), *ECUMICT 2014*,
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_12, © Springer International Publishing Switzerland 2014

processor) on parts of the program that can be done in parallel, then Amdahl's law says that speedup is given by [1]:

$$Speedup = (s + p)/(s + p/N) \quad (1)$$

Working with normalized values, Amdahl speedup formula can be reshaped considering s as percentage of serial processing and p as parallel processing percentage thus $s + p$ can be considered equal to 1 or 100%. The equation (1) can be expressed now like:

$$Speedup = 1/(s + (1 - s)/N) \quad (2)$$

If 50% of the processing instruction are sequential while the rest of them are executed in parallel, for a 20 cores system and 20hours serial single core time, then the overall improvement as speed up is calculated as being 1.9 times compared with serial only situation, i.e. the new execution time is 10.5 hours.

Figure 1 shows the Speed up dependency versus different serial processing fraction for different number of parallel processors, N.

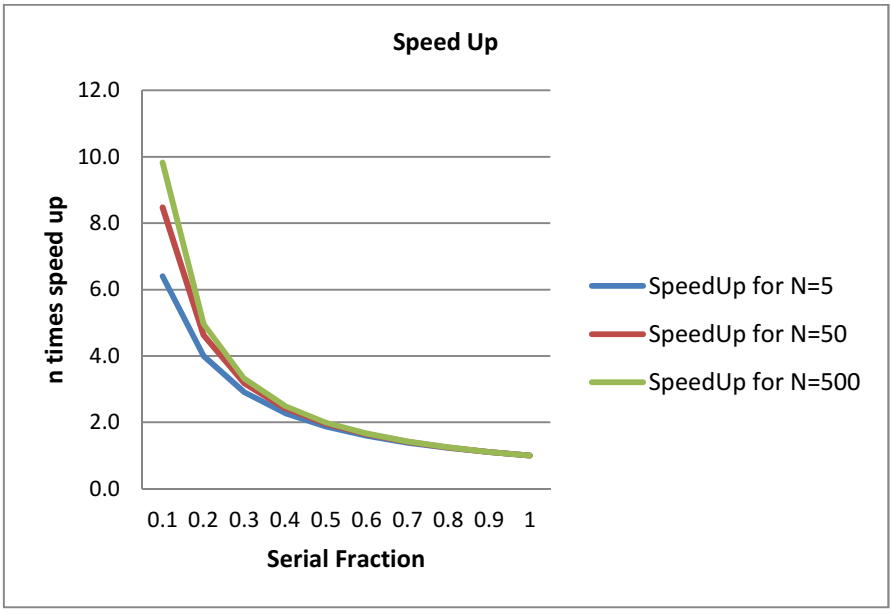


Fig. 1 Speedup improvement under Amdahl's law

2 Delay Considerations

As proposed in [2], for an ideal parallel processor, program execution is divided equally across the number of parallel processors N. For example, if a set of computations must be performed on a data set, then the same computations could be performed in parallel on N processors, each with 1/N of the data set. The time to execute this algorithm is reduced by 1/N, thus the expected delay is:

$$D = \sum_{i=0}^{M-1} \frac{d_i}{N} p(d_i) = \frac{1}{N} \sum_{i=0}^{M-1} d_i p(d_i) \tag{3}$$

In our case d_i is the delay of the *i-th* execution instruction and $p(d_i)$ is the probability distribution for each execution instruction, as expressed in Equation 5:

$$p(d_i) = \frac{Q_i}{\sum_{i=0}^{M-1} Q_i} = G_i \tag{4}$$

where: $Q_i$is the number of instructions with the i^{-th} delay and $\sum_i Q_i$is the total number of instructions executed. This is equivalent to the fraction of instructions G_i executed with the i^{-th} delay.

Therefore our goals are to minimize the expected cost in terms of the delay of the architecture, than to minimize the cost function.

In a realistic (non-ideal) parallel processor, program execution cannot be perfectly divided across the number of N parallel processors. Rather only some portion of the algorithm can be paralleled (the parallel fraction, Fp), while the remaining portion is executed sequentially (the serial fraction, Fs). Thus the total cost J_D is a weighted combination of parallel and serial fractions [2]:

$$J_D = \frac{Fp}{N} \sum_{i=0}^{M-1} Gp_i Dp_i + \frac{Fs}{1} \sum_{i=0}^{M-1} Gs_i Ds_i \tag{5}$$

where:$Fp + Fs = 1$ and $Gp_i$and Gs_i are the instruction distribution fractions with delays: Dp_i and Ds_i for the parallel and serial portions of the algorithm, respectively [2].

A similar cost can be calculated considering the probabilities associated with the data units. Each data unit has a certain parallel or serial delivery probability in the same way as an execution instruction above, Gp_i and s_i. Also, each data unit has a delayed delivery time, basically due to the gap between the data units, which is similar with the delay for processing instruction, Dp_i and Ds_i. They could be different for serial or parallel way of travel.

3 Theoretical Approach

Practice says that if we can make a data transfer with more sessions at a time then the bandwidth of the transferred data will increase. Observing the analogy between the process execution time in Amdahl's Law and the travel time for a data unit we are going to demonstrate the possibility of using the Amdahl's speedup formula for multisession data transfers. More, serial processes considered by Amdahl's Law are similar with basic single stream data transfer over communication channels while multisession communication could be assimilated with parallel processing. In both cases, data processing and data communication, we refer to timing parameters associated with processing time and travelling time respectively. Equivalently, fraction of serial and parallel carried data volumes can be considered instead as in (2) and having the weighted cost as in (5). The number of parallel processes can be equivalent with the number of parallel communication sessions. Accordingly, increasing the number of sessions could affect the data transfer speedup in the same way as speeding up a global process using parallel processing. In the present approach we generate traffic between two stations with a different number of sessions using different constant delay values. Based on data from experiments we can draw some conclusions if they check the law. In practice we have many devices through which data units are transmitted from source to destination. Data units can pass through different topologies and media and errors are affecting transmission delay and transmission quality. In limited error free conditions the Amdahl formula can be used as described below.

Amdahl's Law is basically considering a fraction of sequential computing, let's say F, while the rest of the process, 1-F, is distributed over a number of parallel processors. We are assimilating these with serial communication on a single session which is then improved, in the same way as multiple processors are doing, by adding supplementary parallel sessions. The communication session model is based on the consideration that the data units are traveling in a serial manner, one after another, with a variable data free interval between them. The amount of this inter-units gap is defining the session bandwidth, that being usually manipulated for bandwidth control. Also, this interval is available for the transferof other data units, belonging to other serial sessions, using a TDM like technique (Time Division Multiplexing). Each new extra session can be assimilated with a virtual parallel stream as long as the gap is not completely filled up (channel not congested) and is adding an extra data carry potential to the channel.

Equation (2) shows that the speed up improvement can be evaluated considering a certain percentage of basic serial processing which is similar with the basic serial communication session and the number of parallel processing cores which is similar with parallel communication sessions, out of the basic serial one. The rest of the data volume is equally divided between N parallel sessions.

For our study we replace variables from Amdahl's law, expressed by (1), with the following variables:

s=fraction of data units for a single serial TCP session
N= number of parallel TCP sessions
p=fraction of data units for N parallel TCP sessions

If *s* is the serial fraction of the communication, *BW* is the basic serial bandwidth and *t* is the serial transmission amount of time, then its percentage value is

$$s[\%] = \frac{V_{serial}}{V_{total}} \cdot 100 = \frac{BW_{serial} \cdot \mathrm{t}}{V_{total}} \cdot 100 = k \cdot BW_{serial} \quad (6)$$

V is the data volume. Just For simplicity we can consider 100 data units some of which are included into the serial basic session with a duration of 1 second, generating *k=1* in (6).

It is important to note that, depending on the multiplexing technique, most of the channels are serial type at the physical level while parallel sessions are virtually supported.

4 Test Bench and Results

To test this we used ***iperf*** program. This program is designed to test the internet bandwidth and it is actually based on a packages generator. All tests were done with TCP packets because in this case we have a response confirming that the information came right from one end to another. In the experiments we have been used various delays of the route packets from source to destination. When using UDP protocol, it does not take this into account delays because in this case we do not have ACK packages. Information reaches the other end with some delays (like a video stream) and the results were not relevant. Different delays are impacting on the gap between the data units, so session bandwidth can be manipulated.

In our experiments we run iperf as client on PC1. We use following syntax:
iperf –c 192.168.100.10 –i 1 –t 120 –w 128k –P (number of parallel sessions)
-i 1 is for displaying the results every second;
-t 120 represent the total connections time between iperf client and server;
-w is the TCP window size;
On PC2 we run iperf as server (**iperf –s –i 1**).

The packages travel from PC1 to PC2 through the server. PC1 and PC2 are used to generate packages for bandwidth testing while the server is inserting delays on each packet.

For a preliminary test and to avoid other interferences a simplified test bed was used (figure 2) with systems working under FreeBSD operating system.

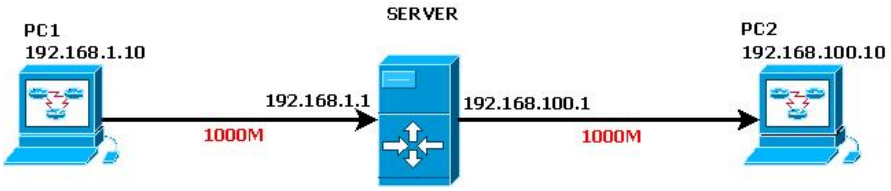


Fig. 2 Test bed representation

The server configuration for passing traffic is presented below:

bge0: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> metric 0 mtu 1500
options=9b<RXCSUM,TXCSUM,VLAN_MTU,VLAN_HWTAGGING,VLAN_HWCSUM>
ether 00:10:18:33:db:48
inet6 fe80::210:18ff:fe33:db48%bge0 prefixlen 64 scopeid 0x3
inet 192.168.1.1 netmask 0xffffff00 broadcast 192.168.1.255
media: Ethernet autoselect (1000baseTX <full-duplex>)
status: active

bge1: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> metric 0 mtu 1500
options=9b<RXCSUM,TXCSUM,VLAN_MTU,VLAN_HWTAGGING,VLAN_HWCSUM>
ether 00:10:18:33:db:49
inet6 fe80::210:18ff:fe33:db49%bge1 prefixlen 64 scopeid 0x4
inet 192.168.100.1 netmask 0xffffff00 broadcast 192.168.100.255
inet 192.168.200.1 netmask 0xffffff00 broadcast 192.168.200.255
media: Ethernet autoselect (1000baseTX <full-duplex>)
status: active

The conducted experiment introduces different values of delays simulating different bandwidth values and respectively different values of the data volume. As described in equation (6) the bandwidth is a direct measure of the serial fraction needed for Amdahl's law.

The experiments use a PC iperf server and an iperf client on the other end. Initial tests with 0 delay mean an ideal traffic condition just to generate a reference value to be considered to identify the limits. Under ideal conditions we got the following result:

[3] 0.0-30.0 sec 2.19 GBytes 626 Mbits/sec

The communication power of a system is not about how it manages to pass traffic through its interfaces but the maximum number of data units that can be processed. We consider this value as the total capacity of the link between the two communicating entities in ideal conditions (delay 0).

In experiments we introduced various delays and values for data units traveling between the two systems. Packet size used was of 128 Kbyte in order to easy calculate how many packets were sent to certain values of traffic. Working with

different sizes and types of data units do not modify the multisession communication evaluation, the differences being just a scaling factor.

If we have a look at the OSI model we see that internet only works up to level 4. In our case for doing this experiment we had to work up to level 7. What we modified in this experiment has to do with the session level. In our case we changed the number of sessions to values under the limit where the systems became unstable due to their limited resources (CPU Pentium M 1.5-1.8 GHz with 512-1024 Mbit RAM).

To expose the results *iperf* program was used and the results are shown in the lines below:

[3] 27.0-28.0 sec 384 KBytes 3.15 Mbits/sec
[3] 28.0-29.0 sec 512 KBytes 4.19 Mbits/sec
[3] 29.0-30.0 sec 384 KBytes 3.15 Mbits/sec
[3] 0.0-30.1 sec 8.38 MBytes 2.34 Mbits/sec

It can be noticed that the transfer is not continuously constant because the involved equipment are working asynchronously.

Different values of packages delays on the server side were used. To apply a specific delay *ipfw* utility was used. This application must be added to the initial configuration of the FreeBSD operating system. To enable this tool, the kernel needs a recompilation with the following options:

options IPFIREWALL
options IPFIREWALL_VERBOSE
options IPFIREWALL_VERBOSE_LIMIT=10000
options IPFIREWALL_DEFAULT_TO_ACCEPT
options DUMMYNET

DUMMYNET option is used to enable shaping.

After recompiling the kernel on the server with the options above we added a set of rules to introduce some delay between the two systems involved in the communication, like in following example:

ipfw add pipe 10 ip from any to any via bge0
ipfw pipe 10 config delay 180

bge0 is the interface name of input packages and "*180*" represents the value of delay, in milliseconds. Different delays of data packets were used for this parameter. The results obtained for a single session have been considered as reference values. All other evaluations we have made are related to these values.

In the following tables the final obtained numerical values were introduced.

Table 1 Experimental data for 1 session

No.	Delay	Sessions	Traffic/s	Packets/s
1	180	1	3.08 Mbits/sec	24.64
2	90	1	7.72 Mbits/sec	61.76
3	45	1	19.5 Mbits/sec	156
3	22	1	41.29 Mbits/sec	330.32

Table 2 Experimental data for 4 sessions

No.	Delay	Sessions	Traffic/s	Packets/s
1	180	4	15.75 Mbits/sec	126
2	90	4	39.8 Mbits/sec	318.4
3	45	4	65.02 Mbits/sec	520.16
3	22	4	103 Mbits/sec	824

Table 3 Experimental data for 8 sessions

No.	Delay	Sessions	Traffic/s	Packets/s
1	180	8	31.15 Mbit/sec	249.2
2	90	8	59.76 Mbit/sec	478.13
3	45	8	84.15 Mbit/sec	673.2
3	22	8	319.5 Mbit/sec	2556

Table 4 Experimental data for 16 sessions

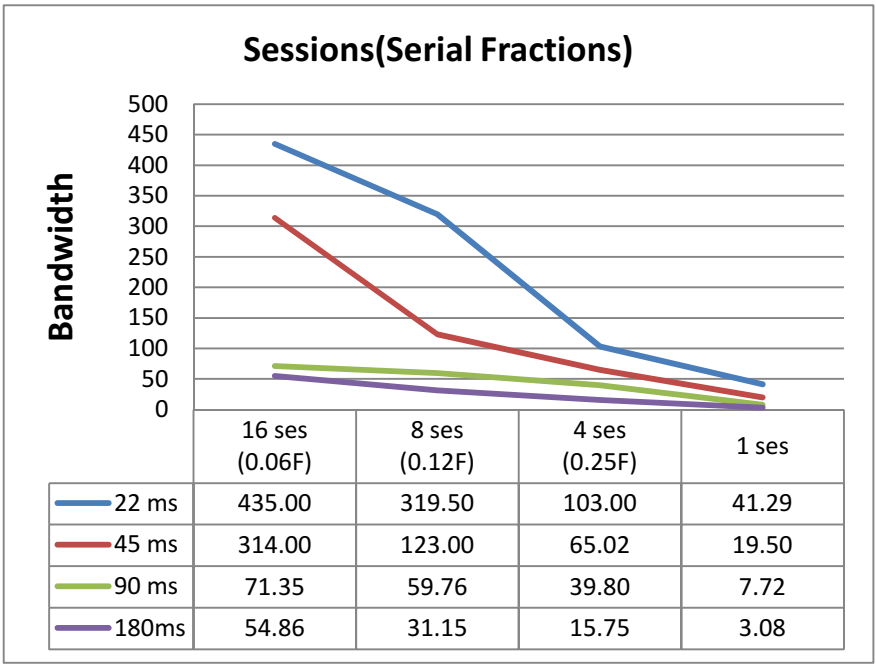
No.	Delay	Sessions	Traffic/s	Packets/s
1	180	16	54.86 Mbit/sec	438.93
2	90	16	71.35 Mbit/sec	570.8
3	45	16	314 Mbit/sec	2512
3	22	16	435.5 Mbit/sec	3484

From the experimental data shown above, the greater the number of sessions means the greater traffic capacity transport between the two systems. For the same package delays and same package size, when using multiple sessions, the traffic is expected to be a multiplied amount of one session. As in the Amdahl's scenario a basic serial fraction has to be always considered (serial session), the difference being distributed within parallel sessions. The experiments were reproduced on different machines with the same results. For q growing number of parallel session the bandwidth increased in a non-linear manner matching the Amdahl scenario (figure 3).

Amdahl law speaks about the serial fractions that can be parallelized. In our study we are considering one session as reference. At four sessions we will have 75% of parallelization meaning 25% of serialization represented by reference

session (we consider reference session as serial). In this case we have a serial fraction of 0.25.

The communication speedup in terms of global bandwidth increases with the number of sessions, as is to be expected because we have both many sessions and more packages. We noted that the increase is not linear and Amdahl law like scenario is now a very pertinent explanation. Different real life factors are contributing to the contamination of the obtained results in a certain manner, the most important contribution being related to the fact that the real delay is not dependent only to explicitly introduced delay but also to systems or interface native delays.



	16 ses (0.06F)	8 ses (0.12F)	4 ses (0.25F)	1 ses
22 ms	435.00	319.50	103.00	41.29
45 ms	314.00	123.00	65.02	19.50
90 ms	71.35	59.76	39.80	7.72
180ms	54.86	31.15	15.75	3.08

Fig. 3 Graphical representation of bandwidth increase vs. serial fraction for different delay values

5 Conclusions

Based on the analogy between the allocated time for serial and parallel processing and the travel time of the data units over serial and parallel data streaming the paper is proposing an extended use of Amdahl's Law in data communication. The communication channels are, in a natural way, serial but multiple streams are possible in the same time as real or virtual parallel sessions. The data units traveling time was introduced in Amdahl's formula to calculate the speedup value in the

same way as the execution time is used in the original application of the formula. Starting from the proposed approach, a new tool for QoS (Quality of Services) evaluation become available in terms of computation of global data rates improvement. The communication efficiency can be optimized by manipulating the way how data are organized into serial and parallel sessions as happen in multisession data services like multicast ones.

As happened in parallel processing, the particular use of Amdahl's Law in communication need some adjustments due to non-ideal behavior of data streams, like interdependency of parallel session or channel saturation. Starting from the Amdahl's model applied in data communication, future researches are expected for describing border behaviors, like congestion phenomena, channel errors effects or other factors involved in disturbing the data flow.

References

1. Amdahl, G.M.: Validity of the single-processor approach to achieving large scale computing capabilities. In: AFIPS Conference Proceedings, Atlantic City, NJ, April 18-20, vol. 30, pp. 483–485. AFIPS Press, Reston (1967)
2. Gustafson, J.L.: Reevaluating Amdahl's law. Common. ACM 31(5), 532–533 (1988), http://doi.acm.org/10.1145/42411.42415, doi:10.1145/42411.42415
3. Cassidy, A.S., Andreou, A.G.: Beyond Amdahl's Law: An Objective Function That Links Multiprocessor Performance Gains to Delay and Energy. IEEE Transactions on Computers 61(8), 1110–1126 (2012), doi:10.1109/TC.2011.169
4. Orzen, S.-N.: Distributed Systems and Artificial Intelligence in Programming. In: 11th International Conference on Development and Application Systems, Suceava, Romania, May 17-19 (2012)
5. Hill, M., Marty, M.: Amdahl's law in the multicore era. IEEE Comput., 41(7), 33–38 (2008)
6. FreeBSD HandBook, http://www.freebsd.org/doc/en_US.ISO8859-1/books/handbook/kernelconfig.html

Fine-Tunning a MAP Error Correction Algorithm for Five-Key Chording Keyboards

Adrian Tarniceriu, Bixio Rimoldi, and Pierre Dillenbourg

Abstract. Different typing devices lead to different typing error patterns. In addition, different persons using the same device have different error patterns. Considering this, we propose and evaluate a spelling algorithm specifically designed for a five-key chording keyboard. It uses the maximum a posteriori probability rule, the probabilities that one character is typed for another, named confusion probabilities, and a dictionary model. Our study shows that the proposed algorithm reduces the substitution error rate from 7.60% to 1.25%. In comparison, MsWord and iSpell reduce the substitution error rates to 3.12% and 3.94%, respectively. The error rate can be further reduced to 1.15% by using individual confusion matrices for each user.

Keywords: error correction, chording keyboard, maximum a posteriori probability, confusion matrix.

1 Introduction

With the technological progress, computing devices have become smaller, portable, or mobile. Due to size limitations, the classic QWERTY keyboard became a suboptimal solution, and was replaced by other methods such as 4×3 multi-tap keypads, mini-QWERTY, or touchscreen keyboards. Though easy to use and efficient, these interfaces are not suitable for certain situations. For example, it is difficult to type a text message while walking. Even so, more than 40% of people do it [15],

Adrian Tarniceriu · Bixio Rimoldi · Pierre Dillenbourg
Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
e-mail: {adrian.tarniceriu,bixio.rimoldi,
pierre.dillenbourg}@epfl.ch

L. De Strycker (ed.), *ECUMICT 2014,*
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_13, © Springer International Publishing Switzerland 2014

which is potentially dangerous as the visual attention is committed to typing and not to the surrounding environment.

Chording keyboards [9] represent a solution for the aforementioned situations. These keyboards allow users to generate a character by simultaneously pressing a combination of keys, similarly to playing a note on a musical instrument. Compared to other devices (such as desktop keyboards, mobile-phone keypads, or touchscreens), they require a smaller number of keys. With five keys, there are 31 combinations in which at least one key is pressed, enough for the 26 letters of the English alphabet and five other characters. If the keys are adequately placed in a position that fits naturally under the fingertips, then we can type with only one hand and without looking at the input device. Therefore, we will be able to use a mobile device even during activities for which vision is partially or entirely committed, like walking in crowded spaces, jogging, or riding a bike.

Besides typing fast and with low error rates, it is important to be able to use a text-entry method without too much training. Previous studies [16, 17] showed that people can learn to type with a five-key chording keyboard in less than 45 minutes. After 350 minutes of practice, the average typing rate was around 20 words per minute (wpm) with a maximum of 31.7 wpm, comparable to iPhone, Twiddler [8] or handwriting. The typing error rate at the end of the study was 2.69%. Being able to automatically correct these mistakes will probably increase the keyboard's ease-of-use and typing speed, because users will not have to stop typing in order to correct errors. In addition, being focused on another activity while typing will probably lead to more errors, so efficient error correction becomes even more important in these situations.

This paper continues our work [18] on an error correction mechanism for chording keyboards. The correction mechanism is based on the maximum a posteriori probability principle (MAP) [5] and for each typed word, it provides a list of possible candidates and chooses the one that is the most likely. Moreover, it takes into consideration the particularities of the text input device. This is motivated by the fact that different devices lead to different error patterns, and knowledge about these patterns can be used to improve the error correction methods. Besides presenting the correction algorithm, we also analyze the factors that influence the algorithm's efficiency.

The paper is organized as follows. Section 2 presents a brief overview of existing text error correction mechanisms. In Sects. 3 and 4, we describe the proposed error correction algorithm and the data set used for evaluation. Section 5 outlines the error correction results. In Sect. 6, we conclude the paper and discuss future research directions.

2 Related Work

A detailed overview of commonly used correction techniques is presented by Kukich in [7]. Research in spelling error detection and correction is grouped into three main categories:

1. Non-word error detection: Groups of *n* letters (*n*-grams) are examined and looked up in a table of statistics. The strings that contain non-existing or highly infrequent *n*-grams are considered errors.
2. Isolated word error correction: Each word is treated individually and considered either correct or incorrect. In the latter case, a list of possible candidates is proposed. These candidates can be provided using several techniques such as minimum edit distance [19], similarity key techniques [10], rule-based techniques [20], *n*-gram techniques [11], probabilistic techniques [4], or neural net techniques [12].
   Most isolated word error correction methods do not correct errors when the wrongly typed word is contained in the dictionary. For example, if *farm* is typed instead of *form*, no error will be detected. Moreover, these methods cannot detect the use of wrongly inflected words (for example, *they is* instead of *they are*).
3. Context dependent error correction: These methods try to overcome the drawbacks of analyzing each word individually by also considering the context. Errors can be detected by parsing the text and identifying incorrect part-of-speech or part-of-sentence *n*-grams [13]. Or, if enough memory and processing power are available, tables of word *n*-grams can be used. Other approaches consider grammatical and inflectional rules, semantical context, and can also identify stylistic errors.

Most of the methods presented above can be applied to any typed text, regardless of the input device. As various input techniques become more and more popular, the classic correction techniques have been improved to consider both the text and the device particularities. Goodman et al. [3] present an algorithm for soft keyboards that combines a language model and the probabilities that the user hits a key outside the boundaries of the desired key. Kristensson and Zhai [6] propose an error correction technique for stylus typing using geometric pattern matching. The T9 text input method for mobile phones can also be included here, as it considers the correspondence between keys and characters to predict words. A strategy that can be applied to chording text input is presented by Sandnes and Huang in [14].

3 Algorithm

Traditionally, text error detection and correction focus on character-level errors, which can be classified into three categories: *deletions*, when a character is omitted; *insertions*, when an additional character is inserted; *substitutions*, when a character is substituted by another character.

The algorithm that we propose is designed only for substitution errors and focuses on individual words, without considering any contextual information. It is based on the maximum a posteriori probability principle, taking into account a dictionary model and the probabilities that one character is typed for another.

For a typed word y, the MAP algorithm will find the string $\hat{x}$, which is the most likely in the sense of maximizing the posterior probability $p(x|y)$ over all $x \in S$. The set S contains all the possible candidate strings. Then,

$$\hat{x} = \arg\max_{x \in S} p(x|y) \tag{1}$$

$$= \arg\max_{x \in S} p(y|x)p(x), \tag{2}$$

where (2) follows from Bayes' rule and takes into account that we maximize with respect to x.

As we focus on substitutions, we can limit the candidate set to words with the same length as the typed word. Considering this and assuming that error events are independent, we can write

$$p(y|x) = \prod_{i=1}^{i=N} p(y_i|x_i), \tag{3}$$

where y_i is the ith letter of the typed word, x_i is the intended letter, and N is the word length. $p(y_i|x_i)$, named confusion probability, is the probability that the character y_i is typed in lieu of x_i. The prior probability, $p(x)$, is given by the frequencies of the dictionary entries in English language. For example, given the typed word $y = oat$ and the candidate $x = bat$, we need to compute

$$F(oat|bat) = p(o|b)p(a|a)p(t|t)p(bat). \tag{4}$$

The set S contains dictionary words with the same length as the typed word. To increase speed, we can use the fact that only a certain fraction of the substitutions occur with non-negligible probability. To describe how this is done, it is useful to represent each character by a five-bit codeword. We choose the first digit to represent the key under the thumb, the second to represent the key under the index, etc. The value of a position is 1 if the corresponding key is pressed and 0 otherwise. So, for instance, 10111, corresponding to the letter b, means that all fingers except the index are pressing the keys (five examples of mappings between key combinations and characters are shown in Fig. 1). In this way, two words can be compared also from a bit distance point of view, as shown in Table 1. Our tests have shown that in 98.5% of the cases, the wrongly typed word differs from the intended one by at most five bits. Hence, for each typed word, we limit the set S to words that differ by at most five bits. Compared to using the edit distance, this method provides less candidates, thus increasing the speed of the algorithm.

In our study, we used the British National Corpus, containing approximately 100 million words [1]. The used dictionary was obtained from this corpus by choosing

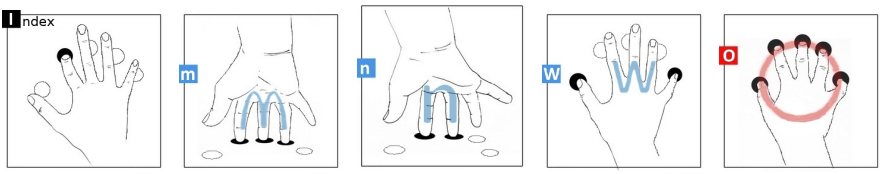


Fig. 1 Examples of letter mappings. "**i**" is given by the initial of the finger pressing the key (index). "**m**" and"**n**" are given by the shape of the fingers pressing the keys. "**w**" is given by the shape of the fingers not pressing the keys. For "**o**", we imagine five dots spread around a circle, and we obtain it by pressing all the keys.

Table 1 Possible candidates for the typed word *oat*

Possible candidate	Binary form	Bit distance
oat	11111 00110 10000	0
bat	11101 00110 10000	1
rat	00010 00110 10000	4

all the items occurring more than five times. It contains 100 944 entries, which include inflected forms such as declensions and conjugations. The prior probabilities were given by the word frequency in the corpus and the confusion probabilities were estimated experimentally.

4 Evaluation Data

In order to gather enough data to evaluate the proposed algorithm, we asked 10 students from our university to type using a chording keyboard prototype. The prototype has the keys placed around a computer mouse and is presented in Fig. 2. We designed the prototype in this way because we wanted the subjects to see a practical application of a chording device: allowing typing and screen navigation at the same time, with only one hand. The buttons are placed so that they can be easily operated while holding the mouse with the palm. The keyboard is designed using an Arduino Pro Mini microcontroller board and communicates with the computer by Bluetooth.

The total amount of data gathered during the experiment consists of 40 345 words, out of which 4052 (10.17%) contain errors. Of these, 3065 (75.64%) are substitution errors. The remaining 987 errors occurred when people did not type a letter (e.g. *hous* instead of *hous**e***), typed an extra letter (*hous**ee*** instead of *hous**e***), the space between words was missing (*thehouse* instead of *the house*), or when whole words were missing, added, or the topic of the sentence changed.

The total number of typed characters is 219 308, from which 5889 are errors. We used these characters to determine the confusion matrix, which is a square matrix with rows and columns labeled with all the characters that can be typed. The value at position *ij* shows the frequency of character *j* being typed when *i* was intended. The

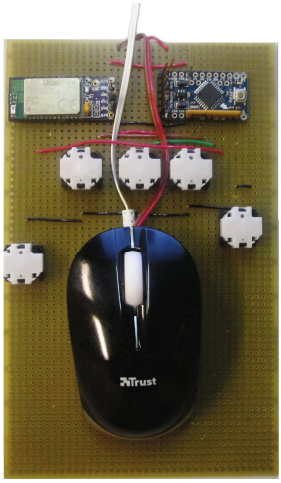


Fig. 2 Chording keyboard prototype used during the typing study

values are given as percentages from the total number of occurrences for character i and represent the confusion probabilities used by the algorithm.

5 Results

The error-correction algorithm was implemented in MATLAB. To evaluate it, we checked the substitution error rates (the number of words containing substitution errors divided by the total number of typed words) before and after applying the algorithm. In the past [18], we compared the results to MsWord and iSpell, to have a reference for the proposed correction method. As this is important for showing the algorithm's error correction ability, we repeat this comparison in Sect. 5.1, with an improvement in the correction method. In Sects. 5.2, 5.3, and 5.4, we show how different dictionaries and confusion matrices affect the correction efficiency, the distribution of errors for different word lengths, and how to improve the correction mechanism using word bigrams.

5.1 Correction Rates

In our previous work [18], we compared the correction results to MsWord and iSpell. The substitution error rate was decreased from 7.60% to 1.59%, which is considerably better than the two references (3.12% for MsWord and 3.94% for iSpell). During that study, the confusion probabilities, $p(y_i|x_i)$, were lower bounded to a fixed value (0.2%), to ensure that any transition between letters is considered. This bound is useful if we have only a small amount of training data, but after

experimenting with larger amounts of data, we noticed that the transitions between certain characters are accurately described by values lower than 0.2%. Considering this, in the following we set the lower bound to be equal to the lowest non-zero confusion probability.

This change in the lower bound reduces the substitution error rate from 1.59% to 1.25%. The correction results (for the proposed algorithm, for the previous work, for MsWord and for iSpell) are shown in Fig. 3, with the substitution error rates depicted by the light bars. The non-substitution errors (dark bars) are not affected by the correction algorithm.

One should not forget that the dictionaries used by the three methods are not the same, and this can affect the results. Moreover, our algorithm is specifically designed for a five key chording keyboard, while MsWord and iSpell can be applied to any text input device with the same results.

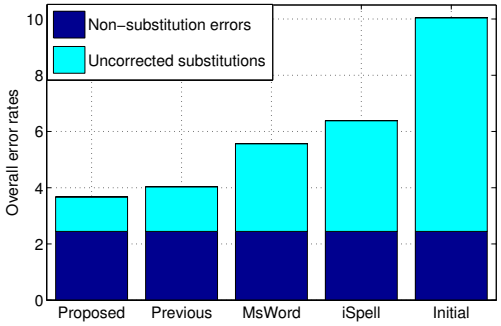


Fig. 3 Overall error rates for the proposed algorithm, previous work, MsWord and iSpell

The MAP algorithm minimizes the error probability by choosing the candidate with the highest posterior probability, which depends on the prior probability and on the confusion matrix. Therefore, by finding better estimates for the prior probabilities and for the confusion matrix, we will be able to further reduce the error rate.

5.2 Dictionary Effect

Besides the error correction rate, we are also interested in the speed of the algorithm. A larger dictionary will most likely reduce the error rates, but the algorithm running time will increase. From a practical point of view, this means slower response times and higher power consumption. Therefore, it is important to find a trade-off between dictionary size and acceptable error rates.

We evaluated the error rate vs. dictionary size for six different dictionaries, denoted as D_2, D_5, D_{20}, D_{50}, D_{100} and D_{500}, respectively. The subscript of the letter

D shows the minimum number of appearances of every dictionary entry in the used corpus. All of them contain inflected words and their sizes are given in Table 2.

Table 2 Dictionary sizes

	D_2	D_5	D_{20}	D_{50}	D_{100}	D_{500}
Size	160 250	100 944	61 364	41 028	29 066	11 288

The results are shown in Fig. 4. As expected, the number of errors decreases as the dictionary size increases, but the relation is not linear. Increasing the size above a certain value has only a marginal effect on the error rate, which may not compensate for the increases in time, processing power, and memory requirements. For example, increasing the dictionary size from 100 944 to 160 250 only reduces the error rate from 1.25% to 1.23%, while increasing the dictionary size from 61 364 to 100 944 reduces the error rate from 1.76% to 1.25%. Finding the optimal dictionary size depends on the desired correction rates and available resources. From a fundamental point of view, reducing the dictionary size by excluding words with low frequency in the corpus is equivalent to setting their prior probabilities to zero. Using less accurate priors can only increase the error probability.

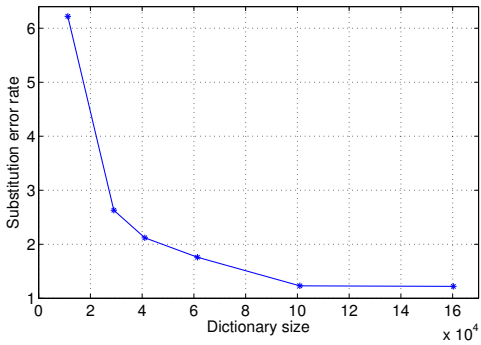


Fig. 4 Substitution error rates for different dictionary sizes

5.3 Confusion Matrix Effect

As already mentioned, the performance of the algorithm depends on the accuracy of the confusion matrix. Given the importance of this matrix, and having in mind that it was determined using the data from all the 10 participants, one might expect that using personalized confusion matrices would lead to better results. To test this, we constructed individual confusion matrices from the text typed by each user. Then, we corrected each user's errors using these matrices. The results were

obtained using 10-fold cross-validation and are given in Fig. 5, where each group of bars represents the substitution error rates for each user, with the individual and with the common matrix, respectively. The horizontal lines are the substitution error rates for the whole typed text, with individual and with common matrices, respectively.

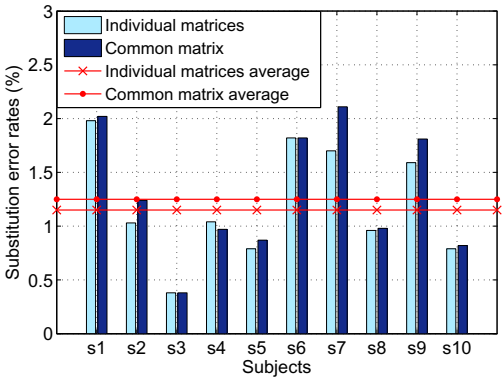


Fig. 5 Post-processing substitution error rates with individual and with common confusion matrices

The use of individual confusion matrices decreased the substitution error rate from 1.25% to 1.15% (this represents a relative reduction of 8%). For 7 of the 10 participants in our typing study the error rates are lower than when using the common matrix. For two participants the error rates are the same, and for only one participant the error rates are higher.

One may also expect that, as people gain typing experience, they will make different types of mistakes. Hence, we built the confusion matrices for each session and used them to correct the errors for that specific session (again, by using 10-fold cross-validation). However, in this case we did not notice any improvement in the error rates.

5.4 Word Length Effect

As shown in the second column of Table 3, the average number of candidates considered by the algorithm depends on the typed word's length, being higher for shorter words. This is explained by remembering that each letter can be written as a sequence of five bits, and that for shorter strings it is more likely to obtain a string with non-zero prior probability by modifying a small number of bits. For instance, for length one strings, i.e. single characters, if we look at candidates that are at a distance of five bits or less, we find all the 26 characters of the alphabet.

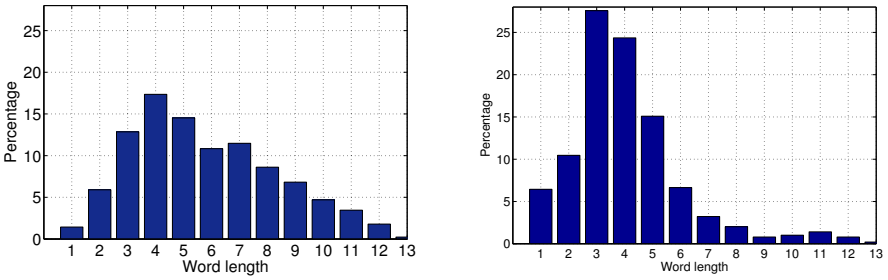
Because there are more candidates for shorter words, the differences between the posterior probabilities of the candidate words are smaller, implying that the error probability is higher. This is confirmed by the third column of Table 3, showing the the ratio between the number of substitution errors that remained after applying the

algorithm, and the number of initial substitution errors, named uncorrected ratio. This ratio is higher for shorter words, meaning that the correction algorithm is less efficient in these cases.

Table 3 Average number of candidates for every encountered word length, and the uncorrected ratio

Word length	1	2	3	4	5	6	7	8	9	10	11	12	13
Candidates	26	345.7	316.4	115.1	97.3	36.3	18.9	4.6	3.2	2.1	1.7	1.9	1
Uncorrected Ratio %	72.7	28.7	34.7	22.7	16.9	9.9	4.6	3.8	1.9	3.5	6.6	7.2	14.3

The fact that shorter words are more difficult to correct is also shown in Figs. 6(a) and 6(b). There, we present the distribution of substitution errors by word length, firstly for all misspelled words, and then after applying the correction algorithm. Notice in Fig. 6(b) that after error correction, shorter words have a much higher contribution to the total number of errors. Indeed, words with length lower than 5 characters count for 69% of the uncorrected errors, but only represent 37% of the initial errors.



(a) **(b)**

Fig. 6 Error distribution for different word lengths. **(a)** Before error correction. **(b)** After error corrrection.

One way to decrease the high error rate among short words is to consider contextual information. For every word with less than 5 letters, we checked its neighbors and considered word bigrams in the correction algorithm. This can be seen as creating new words by concatenating adjacent ones. The length of the new word is the sum of the component words' length, and the prior probability is the bigram probability. This method reduces the substitution error rate from 1.25% to 1.06%, but has the drawback of increased number of computations and the need for a dictionary containing word bigrams and their frequencies.

Another possibility is to analyze part of speech n-grams, as proposed by Church in [2]. For example, if we encounter the one-letter word r, it will probably be an error. The most likely candidates are a (indefinite article) or I (pronoun). By checking

what part of speech is the following word, we can assume with high probability that in the case of a noun the word was *a*, and in the case of a verb it was *I*.

6 Conclusion

In this paper, we have presented an error correction algorithm designed for a five-key chording keyboard. For every typed word, it selects several possible candidates and then returns the most likely one. This is done using the MAP algorithm and the probabilities that one character is typed for another. These probabilities were determined experimentally. Even if the correction algorithm was designed for a specific keyboard and mapping, it can be easily adapted to other input devices by updating the confusion matrix.

The proposed error correction method reduces the substitution error rate from 7.60% to 1.25%, providing a considerable improvement compared to MsWord and iSpell (leading to substitution error rates of 3.12% and 3.94%, respectively). This advantage is due to the MAP algorithm which takes into account the prior distribution of words and the device-dependent confusion probabilities. We also showed that using personalized confusion matrices further reduces the substitution error rate from 1.25% to 1.15%. Using a larger dictionary allows to correct more errors, with the cost of increased search time and memory requirements. Checking word bigrams also decreases the error rate, but again, with the cost of increased complexity.

We have only focused on substitution errors because they represent more than 75% of the total errors. Improving the algorithm by also considering other error types such as missing or extra characters, or words which are not properly separated by white space characters or punctuation signs, will further reduce the overall error rates. Another option for future work is to implement an adaptive approach, starting with a common confusion matrix and updating it based on what one types. Words which are typed more often can have their prior probability increased, becoming more likely than other candidates. One should also be able to add new words to the dictionary.

The comparison between our algorithm, MsWord and iSpell was done by only analyzing the first proposed candidate. We chose this approach because one possible use of the chording keyboards is in dynamical environments, like walking in crowded places or riding a bike, when users cannot continuously look at the typed text. Therefore, the error correction mechanism should run automatically, without requiring user supervision. In more static situations (for example when the keys are placed around a computer mouse, allowing typing and screen navigation with only one device), the most likely candidates can be displayed, and the user will choose the desired one.

References

1. (November 2013), http://www.kilgarriff.co.uk/bnc-readme.html
2. Church, K.W.: A stochastic parts program and noun phrase parser for unrestricted text. In: Proceedings of the Second Conference on Applied Natural Language Processing, ANLC 1988, pp. 136–143. Association for Computational Linguistics, Stroudsburg (1988)
3. Goodman, J., Venolia, G., Steury, K., Parker, C.: Language modeling for soft keyboards. In: Proceedings of the 7th International Conference on Intelligent User Interfaces, IUI 2002, pp. 194–195. ACM, New York (2002)
4. Kahan, S., Pavlidis, T., Baird, H.: On the recognition of printed characters of any font and size. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-9(2), 274–288 (1987)
5. Kay, S.: Fundamentals of statistical signal processing: estimation theory. Prentice-Hall (1993)
6. Kristensson, P.O., Zhai, S.: Relaxing stylus typing precision by geometric pattern matching. In: Proceedings of the 10th International Conference on Intelligent User Interfaces, IUI 2005, pp. 151–158. ACM, New York (2005)
7. Kukich, K.: Techniques for automatically correcting words in text. ACM Comput. Surv. 24, 377–439 (1992)
8. Lyons, K., Starner, T., Plaisted, D., Fusia, J., Lyons, A., Drew, A., Looney, E.W.: Twiddler typing: one-handed chording text entry for mobile phones. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2004, pp. 671–678. ACM, Vienna (2004)
9. Noyes, J.: Chord keyboards. Applied Ergonomics 14(1), 55–59 (1983)
10. Pollock, J.J., Zamora, A.: Automatic spelling correction in scientific and scholarly text. Commun. ACM 27(4), 358–368 (1984)
11. Riseman, E., Hanson, A.: A Contextual Postprocessing System for Error Correction Using Binary n-Grams. IEEE Transactions on Computers 23(5), 480–493 (1974)
12. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning internal representations by error propagation. In: Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1, pp. 318–362. MIT Press, Cambridge (1986)
13. Sampson, G., Garside, R., Leech, G.: The Computational analysis of English: a corpus-based approach. In: Garside, R., Leech, G., Sampson, G. (eds.). Longman, London (1987)
14. Sandnes, F., Huang, Y.P.: Non-intrusive error-correction of text input chords: a language model approach. In: Annual Meeting of the North American Fuzzy Information Processing Society, NAFIPS 2005, pp. 373–378 (2005)
15. Isaac, H., Nickerson, R.C., Tarasewich, P.: Cell phone use in social settings: Preliminary results from a study in the United States and France. In: Decision Sciences Institute Conference (2004)
16. Tarniceriu, A., Dillenbourg, P., Rimoldi, B.: Single-handed typing with minimal eye commitment: A text-entry study. In: The Sixth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, UBICOMM 2012, Barcelona, Spain (2012)
17. Tarniceriu, A., Dillenbourg, P., Rimoldi, B.: The effect of feedback on chord typing. In: The Seventh International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, UBICOMM 2013, Porto, Portugal (2013)

18. Tarniceriu, A., Rimoldi, B., Dillenbourg, P.: Error correction mechanism for five-key chording keyboards. In: The 7th International Conference on Speech Technology and Human-Computer Dialogue, SpeD 2013, Cluj-Napoca, Romania (2013)
19. Wagner, R.A.: Order-n correction for regular languages. Commun. ACM 17(5), 265–268 (1974)
20. Yannakoudakis, E., Fawthrop, D.: The rules of spelling errors. Information Processing & Management 19(2), 87–99 (1983)

Design of a Modular Simulation Environment for Vehicle Mounted Logical Units

Christoph Uran and Helmut Wöllik

Abstract. In spring 2014 the first of a new kind of running event will take place - the *Wings for Life World Run*. It is based on the so-called moving finish line concept where the finish line (represented by a car) gradually catches up with the participating athletes. As a result, the timing of these events has to evolve in order to meet the new requirements. However, developing and especially testing systems that rely on the interaction between different technologies, such as GPS, RFID and mobile communication, can be very costly. Therefore, this paper presents a novel method to simulate the technologies stated above in a parameterizable way. A modular approach to simulate the technologies and emulate their output is illustrated. Furthermore, a software-based prototype capable of simulating a human driver, a car, RFID detections, GPS accuracy and mobile network reliability is presented in this paper.

1 Introduction

Traditional running events are typically limited either by a fixed distance to be completed by the participants (e.g., 100-meter dash, marathon, Olympic triathlon) or by a time limit, where typically the athlete who is able to complete the highest distance wins. However, a new kind of running event is about to emerge - the *Wings for Life World Run* [1] [2]. At this kind of event neither the distance nor the time are determined in advance. Instead, persons who are not able to surpass the required distance at any point of time during the race are counted as finished and can be ranked.

Christoph Uran · Helmut Wöllik
Carinthia University of Applied Sciences, Primoschgasse 8,
A-9020 Klagenfurt am Wörthersee
e-mail: {c.uran,h.woellik}@cuas.at

L. De Strycker (ed.), *ECUMICT 2014,*
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_14, © Springer International Publishing Switzerland 2014

The required distance can be considered a function of the elapsed time. The person that finishes *last* and therefore has completed the highest distance wins the race.

The timing of such events is realized with a *moving finish line*, which is a car that starts a certain time after the athletes have started and catches up with them gradually during the race. This is the so-called *catcher car*. The catcher car has to stick to a predefined speed profile as closely as possible in order to generate comparable results. The slowest athletes will be overtaken first, which eventually leaves the winner, who is the last unranked person.

Sports events are typically timed using RFID (Radio Frequency IDentification) technologies with timing stations on pre-determined points of the track. However, this kind of timing, while perfect for traditional, fixed-distance events, is not suitable for sports events where the time *and* the distance are dynamic. The solution is to mount the timing equipment (including the timing station and its antenna(s)) onto the catcher car.

Another requirement is the ability to organize multiple events of this kind all over the world, at the same time and in a synchronized way. This should result in comparable real-time results from all sub-events across the world. Furthermore, it means that the catcher cars have to be synchronized in terms of their distance traveled at any point of time during the event.

This paper shows how it is possible to simulate the problems that can arise during an event in order to achieve a more cost- and time-efficient testing process of the OBU (On Board Unit) that has to be developed for the catcher car. The problems that should be simulated include unforeseen problems on the track (like collapsed athletes), the driver driving too fast or too slow, unreliable data connections and imprecise positioning. A system overview of the OBU can be seen in Figure 1.

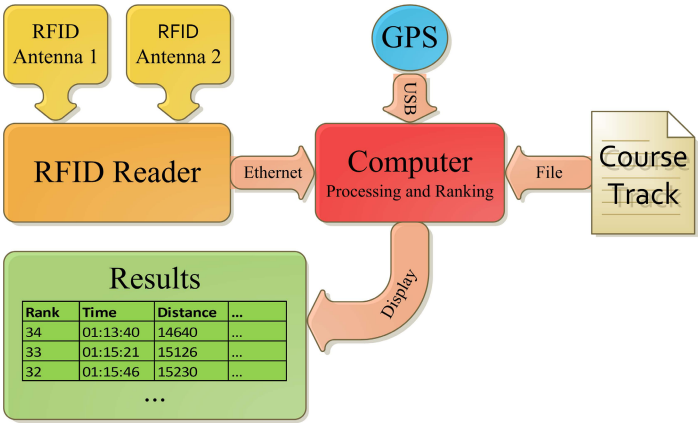


Fig. 1 System overview of the OBU in the catcher car

2 Factors to Be Simulated

In order to understand the developments presented in this paper, it is important to gain an understanding of the factors that are to be simulated. Therefore, this section identifies these parameters and determines their influence for the OBU.

2.1 GPS Accuracy

Generally, GPS (Global Positioning System) is based on the concept of trilateration, which means that it determines the location of the receiving device by measuring its distance to multiple given reference points. In the case of GPS, these reference points are satellites situated in MEO (Medium Earth Orbit) about 20,200 km above sea level. [3]

A robust encoding is used for the signals sent by the GPS satellites in order to compensate the errors that can occur on the distance between the MEO and the receiving device. Therefore, a coded pulse pattern that is comparatively fail-safe is utilized in GPS. This pattern is called PRN (Pseudo Random Noise) because it looks similar to real noise. However, it has information encoded in it. [4, p. 50]

The calculation of the actual position is done by the receiving device. For this, it needs the data sent by the satellites, which includes the exact location and additional information of all available satellites, called the *almanac*. The distance to the satellites is determined by measuring the run-time of the pulsed signals issued by the satellites. The communication is done on the L-band, which utilizes a frequency between one and two GHz, corresponding to a wavelength between 15 and 30 cm. [3]

2.1.1 Identification of Errors

Errors in GPS positioning are induced by constantly changing surroundings. There are no principles to determine the exact impact of the changes. Hence they are random variables within a probability calculation. Most likely random errors come from the electromagnetic wave propagation, electromagnetic noise and uncertainty in the measurement. [3, p. 87]

Thus the resulting run length between the satellite and the receiver can be summarized as shown in Equation 1.

$$\rho_{code} = r + \delta_{eph} + \delta_{iono} + \delta_{trop} - c_{adj}T + \delta_{mp} + v_{rcvr} \tag{1}$$

The range r is the distance between the satellite and the GPS receiver that should actually be calculated from the run length of the signal. The ephemeris error is represented by δ_{eph} and states the difference between the actual satellite position and the satellite position predicted by the ephemeris. The ionospheric and tropospheric

effects are indicated by the δ_{iono} and δ_{trop} respectively. A physical effect that is represented by $c_{adj}T$ is that the waves radiated by the satellite cannot travel at the speed of light in the ionospheric and tropospheric layers of the atmosphere. Therefore, an adjusted c is used. T represents the error caused by the clock of the receiver. Furthermore, the multipath propagation error δ_{mp} and the device dependent receiver measurement noise v_{rcvr} are shown in the equation. Typical magnitudes for these deviations are 3 m for δ_{eph}, 5 m for δ_{iono}, and 2 m for δ_{trop}. $c_{adj}T$ can be reduced to as low as 1 m. DGPS (Differential Global Positioning System) is capable of eliminating local errors such as δ_{eph}, δ_{iono} and δ_{trop}. [5]

According to [6], the height value is the part of the GPS data set that is most severely affected by the weather. Furthermore, the overall accuracy is not very much affected by weather fronts if the GPS receiver is moving. This coincides with the findings of [5] who amount the deviation induced by weather to only 2 m. In [6] it is also stated that "even with surface meteorological data, it is hard, if not impossible, to properly model or predict the wet delay". In this context wet delay describes the delay "caused by water vapor in the troposphere's lower layers". For this reason, the δ_{trop} will not be simulated in the course of this paper, and the value recommended by [5] will be adopted. However, δ_{eph}, T, δ_{mp} and v_{rcvr} will be taken into account in order to be able to modify the simulation for different scenarios (e.g. best-case, standard and worst-case).

2.2 RFID in Sports Events

In many current sports events UHF (Ultra High Frequency) RFID (Radio Frequency IDentification) systems are used to determine the finish and split times of the athletes in a cost-efficient and reliable way. [7]

When comparing RFID to other identification methods such as barcodes, biometry or smart cards some obvious advantages of RFID can be seen. Some of them are the high amount of storable data, good machine readability, no influence of optical covering and comparatively small operating costs. A detailed comparison has been conducted by Finkenzeller in [8, p. 8].

Passive UHF RFID systems such as the one used in this paper utilize modulated backscattering. This is a technology where the reader transmits a modulated signal. The signal is received by the antenna of the chip and thereby transformed into electrical energy, which is used to power a chip inside the transponder (or tag). According to its internal logic the chip changes its own impedance. By doing so, it controls the amount of energy that is reflected by the antenna. The reader then evaluates the reflected signal and is able to determine the code sent by the tag. The small integrated circuits that are used here and the easy production of the integrated antennas makes this technology very cost-effective. [9]

2.2.1 Identification of Errors

In the following section, only the errors induced by the measurement itself are taken into account, but not those that occur during the processing of the measured data.

The most important factor for the result of the read operation is the gap between the reader and the transponder. The transponder has to be supplied with enough power to activate the chip that is responsible for modulating information onto the backscattered signal. In addition the signal must be strong enough to reach the reader and still be processable there. Important influencing factors here are the design of the reader- and transponder-sided antennas and the maximum allowed output power of the reader, which is regulated by various administrations. [8, p. 145]

Another factor is the amount of noise on the used frequency range. Concerning passive UHF RFID systems, the following is stated in [8, pp. 145-146]:

> In backscatter readers the permanently switched on transmitter, which is required for the activation of the transponder, induces a significant amount of additional noise, thereby drastically reducing the sensitivity of the receiver in the reader.

As the timed sports events have more than one participant, there is the possibility that collisions occur when multiple transponders are active at the same time. Multiple-access and anti-collision procedures are used to reduce and recognize collisions. Some examples of multiple-access procedures in RFID are Space Division Multiple Access, Frequency Domain Multiple Access, Time Domain Multiple Access and Code Division Multiple Access. The anti-collision procedures include (Slotted) ALOHA and the Binary search algorithm. [8, pp. 200-219]

2.3 Reliability of Wireless Networks

A permanent data connection between the OBU and a centralized server is necessary during the event in order to ensure a global comparability of the results in real-time. However, an outage may happen at any time during the race. Therefore, multiple wireless communication modules to choose from inside the vehicle are planned. It is the OBU's job to switch between these modules in a transparent way. So it is necessary that the simulation algorithms described in this paper generate outages and other problems just as they could happen in reality.

Arguably the most important factor for the reliability of wireless networks is the mobile signal variation, which happens due to three reasons:

Physical path loss: Occurs during the free space propagation. The signal strength decreases exponentially with the distance.

Large scale or slow fading: Describes the attenuation due to objects that are blocking the direct path between the sender and the receiver. Due to the diffraction of radio waves, not all the area behind the object is affected. Depending on the wavelength of the signal and the shape of the object, the signal can be

received again in some distance behind the object even without a direct line of sight.

Multipath or fast fading: Influences the signal by convoluting or erasing parts of the signal due to the multipath propagation.

Further quality influences include the properties and conditions of the surroundings, electromagnetic noise and the manufacturing quality of the wireless modules.

2.4 Computer-Directed Human Behavior

It is not common that humans are instructed by computers how they should drive a vehicle in terms of the exact velocity. In a usual environment the opposite is the case. However, this is exactly what the situation at hand is about. Because of the uncommonness of this combination, there are no reliable sources in terms of how humans respond to the instructions provided by the DAU (Driver Assistance Unit). Therefore, it was necessary to make one important assumption according to one's own experiences and interviews with other persons.

The assumption is that the driver checks the DAU more often if the current speed differs greatly from the speed instructed by the DAU. This is because a driver tries to decrease the difference as fast as possible. One can compare this behavior with that of a driver getting driving directions from a navigation system. When following a highway for a long time, the navigation system is only consulted sporadically. However, when navigating to a destination within a city and on a route with many branchings and crossroads, the driver will often check the navigation system.

3 Methodology

This section provides a structural overview of the most fundamental parts of the developed system, how they work and how they interact with each other. Before the actual development process started, an evaluation of existing simulation modules has been conducted. The result of this evaluation was not satisfying, because there are no simulation solutions available, that cover such a broad spectrum, as it is intended here.

3.1 Overall Concept

Before starting the development of the modular simulation unit it was crucial to specify the components and the interfaces between each other. To achieve the required modularity, the following design decisions were made in advance:

- Fragmentation of different simulation tasks
- Parameterizability through configuration file(s)
- Development of a simulation controller, providing the modules with instructions
- Flexible interfaces for the communication between the simulation modules

The next step was to define how the simulation modules, the controller and the external devices, which the simulation is made for, should interact with each other. Figure 2 shows this in a comprehensible way.

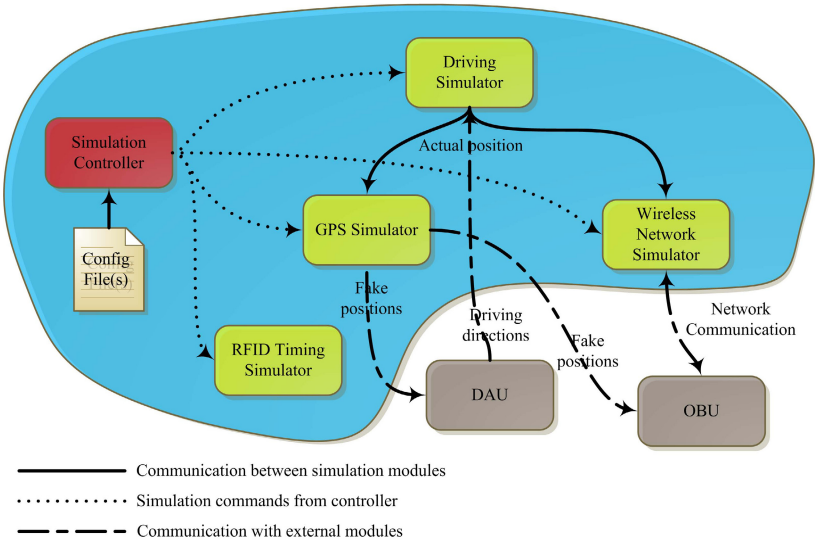


Fig. 2 Overall concept and interfaces of the simulation unit

The most important module of the overall simulation is the simulation controller, displayed in the upper left corner of Figure 2. It provides the other simulation modules with commands that parameterize their simulation tasks. The controller depends on the information in the configuration file which is created by the user. It contains the general information about the track and the surroundings which are to be simulated, as the following list shows:

Course definition: A list of GPS coordinates representing the track.
GPS accuracy: Definition of the severity of parameters influencing the GPS accuracy.
Driver accuracy: Description of how accurate the driver is able to follow the instructions given by the DAU.
RFID detections: This parameter defines the detection frequency.
Wireless network throughput: A representation of the throughput in up- and download on any given position on the track.

All the simulation modules are provided with simulation commands from the controller, whose task it is to decide about special events triggered by certain actions. Also the simulation modules communicate with each other. This is necessary because they have to supply each other with information. For example, the GPS simulator and the wireless network simulator need to retrieve the current location from the driving simulator.

3.2 Algorithm Design

This section describes the algorithms designed for the simulation environment. The key functionality of each algorithm is shown below.

3.2.1 Driving Simulator

In Figure 2 the driving simulator was shown as a single block. However, due to modularity reasons and to make the simulation more realistic, the decision was made to split this module into two separate parts. These parts are:

The DriverThread: This sub-module simulates the behavior of the driver, including the intervals the driver checks the DAU, the driver's use of the throttle/brake and the overall accuracy of the driver. All these values are randomized within the degree of freedom specified in the configuration files. This is especially important for the simulation of human behavior.

The CarThread: It simulates the more complex and low-level parts of the overall module. Some of them are the acceleration and braking pattern of the car and the calculation of the simulated distance and position on the intended track. To facilitate the simulation, linear acceleration and braking patterns have been assumed. Furthermore, possible turns on the track do not influence the speed of the car in the simulation. The calculation of the current distance $dist_{cur}$ can be achieved by adding the distance traveled since the last calculation to the previous distance $dist_{prev}$. Due to the linear acceleration pattern, the distance traveled since the last calculation can be computed as shown in Equation 2. The calculation of the current GPS coordinates is especially important, because they have to be forwarded to the GPS simulator and the wireless network simulator. Equation 3 shows an exemplary calculation of the current latitude lat_{cur}. $dist_{cur}$ represents the distance currently traveled, as calculated in Equation 2. The latitudes/distances of the current and next position in the track definition are represented by lat_1, lat_2, $dist_1$ and $dist_2$. The result is the latitude of a given point with known distance between two points with known latitudes and distances, whereas a straight line between the known points is assumed. The calculation of the current longitude builds upon the same principle.

$$dist_{cur} = dist_{prev} + \left(\frac{v_{prev} + v_{cur}}{2} \right) * t \tag{2}$$

$$lat_{cur} = lat_1 + (lat_2 - lat_1) \frac{dist_{cur} - dist_1}{dist_2 - dist_1} \tag{3}$$

3.2.2 GPS Simulator

As could be seen in Section 2.1, the accuracy classification is a very complex task. Even with exact data concerning the surroundings, weather and the satellites, it would be nearly impossible to simulate the results of the positioning process. Furthermore, the influences of the receiver's quality would be very hard to numeralize. The parameterization by the user would also require a thorough understanding of weather influences, multipath propagation and the technologies used in the GPS receiver from the user.

Therefore, the decision was made to let the user specify the maximum error according to the desired situation to be simulated. This maximum error is then weighted with a normally distributed random value (with a mean of zero and a standard deviation of one) and appropriately applied to the latitude and longitude values.

3.2.3 Wireless Network Simulator

When the catcher car drives on a track, the communication unit will experience rising and falling reception levels and a transition between different mobile communication technologies with different properties. This also includes complete outages if they are specified in the network coverage file provided via the simulation controller. For this simulation the attenuation is modeled with a cubic function, which is similar to the real behavior.

The network coverage file specifies zones that are covered by a certain base station and its properties. These properties are start distance, end distance, maximum upload/download speed and upload/download speed at the margin of the cell. If the current distance is not covered by any cell, the upload/download speed is zero.

According to the distance, a cubic attenuation factor f can be calculated, which is between 0 (no attenuation) and 1 (full attenuation). If the maximum bandwidth B_{max} and the marginal bandwidth B_{mar} are known from the network coverage file, the current bandwidth B_{cur} can be calculated as shown in Equation 4.

$$B = B_{max} - f \times (B_{max} - B_{mar}) \tag{4}$$

3.2.4 RFID Timing Simulator

The RFID Timing Simulator simulates the athletes passing an RFID timing station. However, one problem for the simulation of the passings is that in a real sports event they can be distributed in a broad variety of ways. Furthermore, the simulation is required to be parameterizable, which means that the user should at least be able to configure when the main wave of athletes should pass. For this reason, a wide range of distributions and functions was evaluated and it was researched how these can be used in the simulation, but still be parameterizable. However, most of these distributions pose the problem that they cannot guarantee that all athletes have passed the timing point at any given point of time. Most notable for this is the Chi-squared distribution and the hyperbolic tangent function.

In the end the decision was made to use linear ascends and descends, which are staggered into parts. Each of these parts represents a discrete amount of time, whereas the magnitude of the part represents the relative number of athletes passing the timing point. The position and magnitude of the part where most of the passings occur (also known as the main wave) can be configured by the user. Furthermore, a randomization factor has been inserted in order to model a real sports event.

4 Results

After the algorithm design, a prototype of the software was implemented in Java. This software provides the necessary functionalities to let the user specify all the important parameters of the event to be simulated and execute the simulation in real-time. Some important parameters are a list of transponders, a specification of the track and its mobile coverage, the behavior of the driver and the car, the overall accuracy of the GPS receiver and the maximum upload and download speed of the wireless communication module. Every detected athlete can be seen by the user.

Figure 3 displays the behavior of four important output parameters over the driven distance. The first graph shows the progression of the car's speed during the event. A target speed of 8 km/h (2.22 m/s), which is marked by the dashed reference line has been set for this simulation. Furthermore, a deliberately low accuracy of the driver has been chosen, which explains the high fluctuations of the speed. The cumulative number of finished athletes is shown in the second graph. According to the parameterization, the main wave of athletes is passed at around 70 percent of the track. In the last graph, the progression of the up- and download speeds can be seen. The simplified model that is applied here, shows that both speeds get lower and the fluctuations gets higher with a higher distance from the center of the cell.

Another important type of data to be simulated are the GPS coordinates extracted from GPS modules. Figure 4 shows a histogram of the deviations between the "real" GPS coordinates according to the track specification and the simulated coordinates calculated by the GPS simulator. For this simulation, a maximum deviation of five

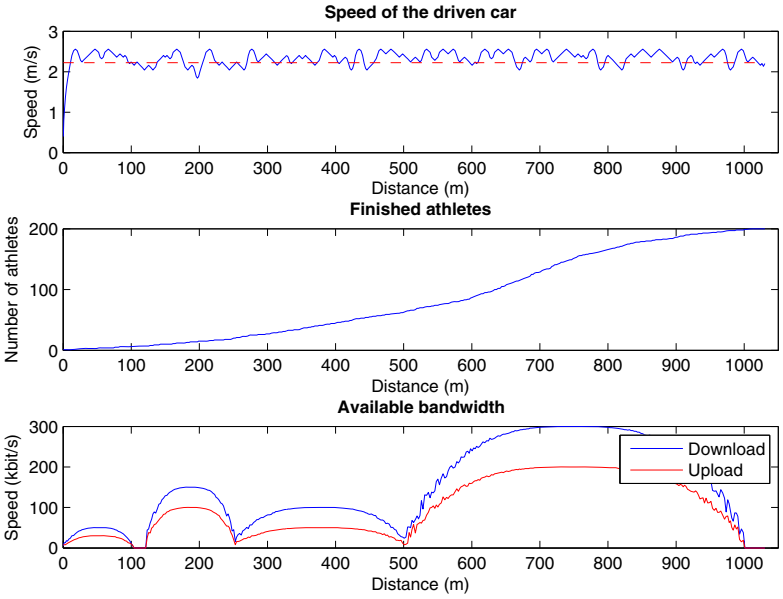


Fig. 3 Graphically formatted results of a simulation

meters has been specified (see 3.2.2), which matches exactly with the simulation results seen in the figure.

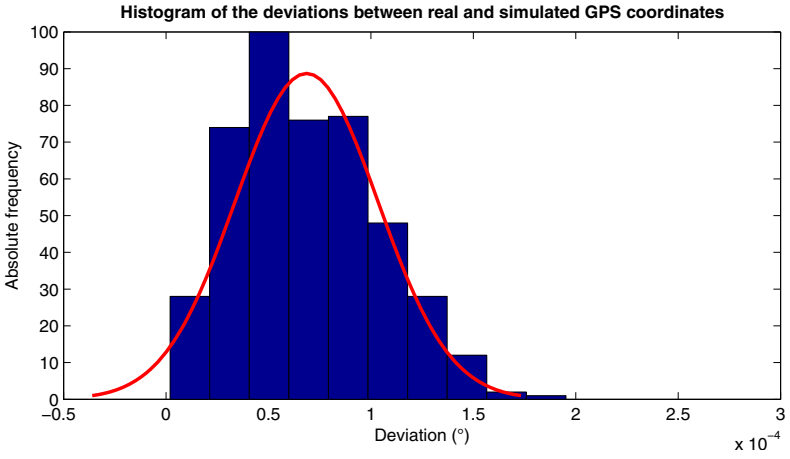


Fig. 4 Deviation of the GPS coordinates

In the current version some features are still missing. This includes the ability to simulate unexpected events such as collapsed athletes in front of the car or mobile communication outages caused by the mobile communication unit. Furthermore, the ability to employ statistical distributions for the detection rates of the athletes has not been implemented yet. However, this would make the simulation more realistic.

Altogether, the system provides a good foundation for further development in the area of simulation environments for vehicle mounted logical units.

5 Conclusion

After conducting all the necessary research and implementations, it can be said that, when testing vehicle mounted logical units, a simulated approach holds significant advantages over conventional testing. The main reasons for this are of organizational and financial nature, but also the parameterizability of the simulation plays an important role. Already slight modifications of the developed prototype result in a working test unit for the OBU as soon as its first version has been developed.

The most significant scientific conclusion of this paper is the general model of a modular and parameterizable simulation unit shown in Section 3.1. Due to the possibility to adapt this model for other simulation tasks and to re-use the existing simulation modules including their algorithms and implementations, a rapid creation of simulations for other vehicle mounted logical units is possible.

It is suggested that the next steps for the further development of this simulation environment are to specify the hardware and the interfaces in order to enable the simulation to communicate with the OBU and the DAU as soon as they are developed. Furthermore, a mathematical thesis could deal with the evaluation and implementation of adequate distributions for the detection sequences of the transponders. Also a thesis in the area of network modeling or radio frequency communication could create and implement a more realistic model for the properties of the data channel between the OBU and the central server.

References

1. Woellik, H.: Sports timing with moving RFID finish line (2013)
2. Wings for Life: The World runs as one (2013), http://www.wingsforlifeworldrun.com/en/news/article/the-world-runs-as-one-28/
3. Mansfeld, W.: Satellitenortung und Navigation. Vieweg+Teubner (2010)
4. Kaplan, E.D., Hegarty, C.J.: Understanding GPS. Artec House, Inc. (2006)
5. Drawil, N.M., Amar, H.M., Basir, O.A.: GPS localization accuracy classification: A context-based approach. IEEE Transactions on Intelligent Transportation Systems (2013)
6. Gregorius, T., Blewitt, G.: The effect of weather on GPS measurements. GPS World (1998)

7. Howell, D.R.: Timing tag. US Patent App. 12/553,369 (2009)
8. Finkenzeller, K.: RFID Handbook. Carl Hanser Verlag (2003)
9. Nikitin, P.V., Rao, K.V.S.: Antennas and propagation in UHF RFID systems. In: IEEE International Conference on RFID (2008)

Out-of-Band Password Based Authentication towards Web Services

Jan Vossaert, Jorn Lapon, and Vincent Naessens

Abstract. A username/password combination is still the most commonly used method for user authentication in a Web based context. Users are familiar with this type of authentication and the registration phase for new users is straightforward. It, however, also has several disadvantages. For instance, users have to deal with an explosion of different usernames and passwords. This may cause users to use short easy to remember passwords, use the same password for multiple services, etc. Further, if malware is running on the workstation, it can eavesdrop on the username and password when entered via the keyboard. Therefore, this paper presents a solution that maintains the familiar wide spread password based authentication mechanism but tackles both the password management problem and prevents malware running on the workstation from stealing the user's credentials. The usernames and corresponding passwords of the user are stored encrypted on his mobile device. The mobile device handles the authentication towards the service provider and transfers the established authenticated session to the workstation. Subsequently, the user can further consume the service on the workstation without having to enter his credentials on the workstation.

Jan Vossaert · Jorn Lapon · Vincent Naessens
KU Leuven, Department of Computer Science, Technology Campus Ghent,
Gebroeders De Smetstraat 1, 9000 Ghent, Belgium
e-mail: {Jan.Vossaert,Jorn.Lapon,Vincent.Naessens}@cs.kuleuven.be

L. De Strycker (ed.), *ECUMICT 2014,*
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_15, © Springer International Publishing Switzerland 2014

1 Introduction

Many online service providers offer personalized services to users. This requires users to login. A username/password combination is still the most commonly used type of credential for authentication in a Web based context. With the increasing amount of online services, users have to remember more and more passwords. User circumvent this problem by reusing passwords, by writing them down or by storing them in a file on the workstation. This, however, introduces several security and usability issues. Moreover, even if users would use unique strong passwords for each service provider and remember them by heart, malware running on the workstation is still capable of eavesdropping on the credentials when they are entered via the keyboard. This is especially an issue if a public/untrusted workstation is used.

Several existing solutions try to tackle one or more of these issues. Many browsers allow the user to store their credentials in a password protected vault. The user can unlock the vault and use the contained credentials by entering the master password. While this provides more protection than storing the credentials in a plaintext file, the user's credentials can still be stolen by malware eavesdropping on the entered master password. Moreover, this solution limits mobility of the user as the credential vault is typically only available on devices owned by the user. Other solutions use two-factor authentication. These systems require the user to know the correct username/password combination and to prove ownership of a specific hardware device, typically a SIM card/mobile phone. This system ensures that knowledge of the username/password alone is not sufficient to gain access to the personalized service. It, however, does not prevent usernames and passwords from being stolen by malware, it requires modifications at the service provider to support the system and the user typically has to release his mobile phone number to the service provider.

This paper presents a new solution for password management. The credentials are stored on the mobile device of the user. This ensures that users can choose strong passwords since they no longer need to be remembered by heart and it ensures portability as the user typically has his mobile device with him. When the user wants to access a restricted service via a workstation, the browser delegates the authentication to the mobile device. The mobile device authenticates the user towards the remote service provider using the stored username and password. The received session cookies are then transferred to the workstation. This allows the user to access the service via the browser on the workstation without exposing his credentials to the workstation.

The rest of this paper is structured as follows. Section 2 points to related work. The design is presented is section 3, followed by a discussion of the prototype in section 4. Subsequently, the design is evaluated in section 5 and, finally, conclusions are drawn in section 6.

2 Related Work

Two-factor authentication solutions [1, 2] require users to prove knowledge of a valid username/password combination and possession of a mobile device. This prevents that a stolen username/password combination alone is sufficient to impersonate the user. The system, however, requires that service providers support two-factor authentication. Currently, most service providers still only require a valid username/password combination.

LastPass [3] is a plugin available for common browser such as Google Chrome, Firefox and Internet Explorer. The plugin stores the username/password credentials of the user for Web services on the hard disk of the workstation. The passwords are encrypted using a key generated from a master password. LastPass automatically fills in the username and password in the designated fields of the HTML page. This systems protects the credentials of the user by storing them encrypted on the hard disk. However, if a key logger registers the master password, all credentials stored in the vault can be compromised.

PwdHash [4] is a browser extension that transparently generates a different password for each service provider based on a master password entered by the user, some unique data associated with the service provider and a private salt stored on the client machine. The password is generated by applying a cryptographic hash on a combination of these sets of data. The user, hence, only needs to remember one password. Malware running on the workstation can, however, still intercept the master password from which all other passwords are derived.

Other solutions use cryptographic tokens to replace username/password authentication. The solutions presented in [5, 6] rely on a smartphone that manages a strong authentication token. In PorKI [5], the smartphone transfers a proxy credential to the browser on the workstation. The browser can then use this credential to set up mutually authenticated HTTPS sessions with servers. Since the proxy credentials have a limited lifetime, the credentials can only be abused by malware on the workstation for a limited vulnerability interval. In SmartIDM [6], the smartphone authenticates the user towards the service provider. The service provider can link the authentication via the smartphone with the session initiated in the browser. The user, therefore, no longer needs to transfer his authentication credentials to the workstation. Both solutions, however, require modifications to the standard password based authentication logic of the service provider. Another popular solution is hardware tokens, such as RSA SecurID devices. They offer high-grade security and typically do not require any changes to the users workstation. However, hardware tokens are often issued for a single application, which can potentially give rise to a usability issue. In addition, similar to many smart card strategies, issuing these tokens brings about an extra hardware cost per new service provider.

Another approach is graphical passwords [7, 8, 9] During registration, the user clicks certain image areas or orders images in a certain order. This pattern is then used as a secret to authenticate the user. This approach assumes it is much easier for humans to remember these patterns than a strong text password. Each image can trigger the user to remember the unique secret pattern. A click pattern can,

however, also be intercepted by malware and it requires significant modifications of the service provider's authentication logic.

3 Design

This section first lists the requirements of the system, followed by a general description of the system. Finally, the protocol is discussed in more detail.

3.1 Requirements

R_1 User can complete username/password based authentication procedures of Web based service providers without having to transfer their credentials to the workstation.
R_2 No modifications of the service provider's password based authentication logic are required.
R_3 The solution does not require authentication to be bound to a particular workstation.

3.2 General Approach

The user (*U*) carries a mobile device (*M*). The mobile device contains the usernames and corresponding passwords of the user. For each service provider (*SP*) with which the user is known, a different username and password is stored. The workstation is used to access a remote Web service. The service provider requires the user to go through a password based authentication procedure before granting access to its services. Instead of entering the required username and password via the workstation, the authentication is delegated to the mobile device. The mobile device uses the stored usernames and passwords to authenticate the user towards the service provider. The service provider checks the credentials and, if the verification is successful, generates one or more session cookies and sets the session as authenticated. The mobile device can now forward the session cookies to the browser (*B*) where they are used in subsequent requests to the service provider. When the browser refreshes the page, the user can, hence, access the protected resources of the service provider.

To transfer the session cookies from the mobile device to the browser, a Web based storage server (*S*) is used. This setup is used since the workstation and mobile device are not necessarily connected to the same local network (e.g. public workstations in hotels or libraries). Hence, direct communication between these two devices might not be possible. This, however, requires the mobile device to have Internet

access. The user is registered with the storage provider, the authentication credentials are stored on the mobile device.

3.3 Protocols

Figure 1 presents the protocol for password based authentication towards a remote service provider. We assume the user has opened his mobile password management application and entered his password to decrypt the password storage. The application is, hence, ready to be used for authentication.

First, the user requests access via the Web browser on the workstation to a protected resource of the service provider (1). The service provider responds with an HTML page containing input fields to enter the user's username and password (2). Instead of having the user enter his credentials on the workstation, the browser generates a QR code[1] containing a unique randomly generated identifier (*id*), the URL of the login page (url_{sp}) and the public key (pk_b) of a generated key pair (3-4). The user uses his mobile device to scan the QR code with the password management application. Subsequently, the application retrieves the user's username and password for the specified service provider from local storage (5). The mobile application now requests the login page from the service provider (6). It hereby receives the initial session cookies and extracts the fields in which the username and password must be entered, any hidden fields contained in the page and the URL to which the authentication must be posted from the login page (7). The mobile device now submits the username and password, together with the hidden fields from the login page to the service provider via an HTTP POST (8). The service provider verifies the authentication and, if the verification was successful, returns a personalized Web page to the mobile application, together with one or more session cookies (9). The mobile application subsequently encrypts the received session cookies with the public key received from the browser (10). The smartphone now transfers the encrypted session cookies to a remote Web based storage service where they are temporarily retained under the unique identifier *id* (11). The cookies are transferred over an HTTPS connection with both server and client authentication. Client authentication is performed with the key received by the user during registration with the storage provider. After showing the QR code, the browser started polling the remote storage server to obtain the encrypted cookies retained under the unique identifier *id* (12). No client authentication is required for requesting the encrypted cookies, the correct identifier is sufficient. Once the cookies are uploaded by the mobile device, the storage service returns the encrypted cookies and removes them from memory (13). The browser now decrypts the encrypted session cookies with its private key and starts using them during communication with the service provider (14-15). When the browser now launches a new request for the protected resource, the service provider grants access based on the used session cookies (16-17).

[1] A Quick Response (QR) code is a two-dimensional barcode that can contain up to a several hundred bytes of data.

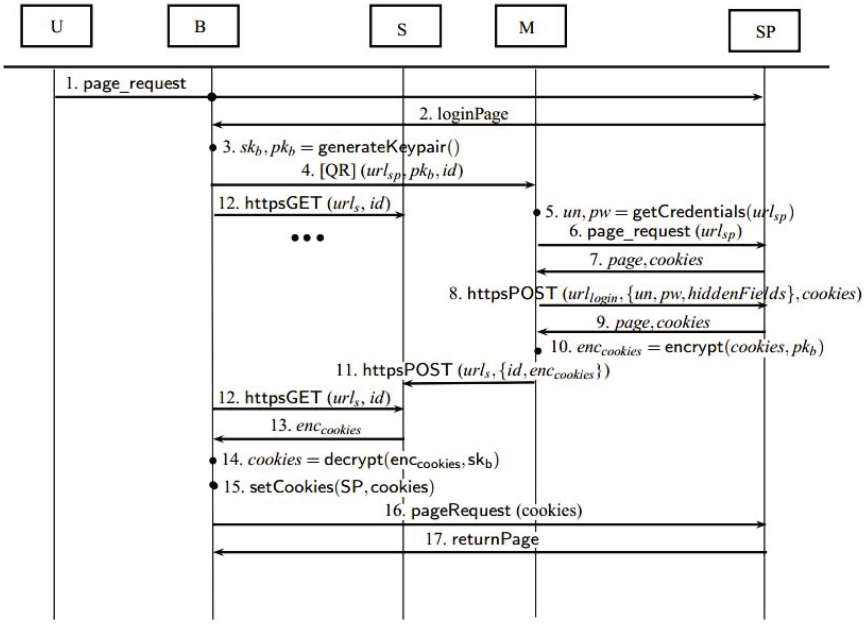


Fig. 1 The authentication protocol

4 Realization

For the functionality of the mobile device, an Android app has been developed. The Android application stores all usernames and passwords of the user encrypted with an AES 128 bit key. This key is generated based on a password that the user enters when starting the application. The cryptographic key is generated using the `PKCS#5`[2] standard.

The browser Firefox has been extended with the required additional functionality via a browser plugin. The plugin is developed using the Firefox Add-on Software Development Kit (SDK) version 1.14. Part of the functionality is also realized by adding a Java applet. The applet generates a 2048 bit[3] RSA keypair for securing the communication between the browser and the mobile device and an 128 bit unique identifier for referencing the cookies. It uses the ZXing library[4] to generate and show the QR code in the browser. Finally, the applet implements the polling of

[2] The `PKCS#5` standard is part of the Public-Key Cryptography Standards (PKCS) series and provides recommendations for the implementation of password-based cryptography.

[3] A 2048 bit key offers a good tradeoff between performance and security. An overview of multiple security recommendations regarding key length can be found on `http://www.keylength.com`

[4] More information about the ZXing library can be found here: `http://code.google.com/p/zxing/`

the storage server. Once the cookies are obtained, the plugin sets the cookies in the browser session with the service provider. The user can start the authentication protocol by clicking on a widget injected in the login HTML page of the service provider by the plugin. The plugin searches the Web page for input fields of type *password* so that it can inject the plugin activation handle on a convenient place in the HTML page (i.e. next to the username/password input fields). A similar strategy can be used to support other browsers.

The functionality of the storage server is implemented using Java servlets. The cookies are temporarily stored using the key value store Memcached[5]. The storage service uses HTTPS client/server authentication to prevent unauthorized users from flooding the service with data. For retrieving the encrypted cookies, knowledge of the unique identifier is sufficient.

The prototype has been tested, and found compatible, with several popular Web based service providers such as Facebook.

Two extensions will be realized in future work. First, some service providers send redirects (i.e. HTTP 301 status code) during the authentication process. The middleware currently does not follow these redirects which results in a failed authentication. For future work, support will be added by interpreting the redirect response and resubmitting the user's credentials to the new page. Second, some service providers use JavaScript generated login pages. Support will be added for correctly parsing these pages.

5 Evaluation

5.1 Requirements Evaluation

The mobile device authenticates the user towards the service provider. The received session cookies are transferred to the workstation. This allows the user to establish an authenticated session with the service provider on the workstation without transferring his login credentials to the workstation. This satisfies requirement R_1.

The prototype has been successfully tested with several popular service providers, satisfying requirement R_2.

For the user to be able to use a particular workstation, it is only required to have an Internet connection and a browser extended with the plugin. Since plugins can typically be installed without root privileges, public workstations can be equipped with the required software as needed. This satisfies requirement R_3.

[5] More information about Memcached can be found here:
`https://code.google.com/p/memcached/`

5.2 Security Evaluation

5.2.1 Attacker Model

We assume attackers can monitor and manipulate network communication between the devices in the system. Further, the software running on the workstation and storage server can be compromised by malware. We assume that the user has an unrooted smartphone. Regarding the cryptographic capabilities of the attacker, the Dolev-Yao attacker model [10] is used. Attackers can not break cryptographic primitives, but they can perform protocol-level attacks.

5.2.2 Security Evaluation

The visual communication channel used between the workstation and the mobile device to transfer the public key of the workstation prevents man-in-the-middle attacks. The close range over which the data is transferred makes it very hard for an attacker to insert or modify the data without the user noticing. The mobile device is, hence, assured that only the workstation has the corresponding private key and can decrypt the session cookies.

An attacker that eavesdrops on the visual communication between the workstation and the mobile device can obtain the public key of the workstation, the unique identifier and the URL of the service provider. The authentication mechanism of the storage server ensures that only registered users can submit encrypted cookies. An attacker can, however, retrieve the encrypted session cookies from the storage server. They are, however, encrypted with a key only known to the workstation. Hence, even a malicious storage provider cannot access the session cookies of the user.

The passwords should be submitted by the mobile device to the service provider over an HTTPS connection. This prevents eavesdroppers from obtaining the username and password of the user. However, some service providers only support HTTP communication. The mobile device notifies the user if his credentials are transferred over an unprotected channel.

Although malware running on the workstation cannot obtain the username and password, the session cookies can be used to access the services of the service provider. The solution presented in this paper protects the long term credentials of the user (i.e. usernames and passwords) by transferring a short term token (i.e. session cookie(s)) to the workstation. This results in a much shorter vulnerability interval.

Since the username and password no longer need to be remembered by heart, the user can choose unique strong, possibly even computer generated, credentials. This impedes brute force attacks on the credentials of the user or attacks relying on reuse of passwords.

The credentials of the user are stored on the smartphone of the user. The security relies on the sandbox model of Android where applications cannot access data stored

by other applications. Moreover, the credentials are encrypted using a key stored in the *keychain* of the device. Since only the application that stored the key in the *keychain* can access the specific key, application level malware cannot be used to obtain the encryption credentials of the user. Moreover, the key is only unlocked if the user unlocked the device (e.g. by entering the PIN or swipe pattern), protecting the usernames and passwords if the smartphone is lost or stolen. While the system protects again application level malware, malware that compromised the OS and acquired root privileges can obtain the user's credentials. However, the smartphone is typically a device controlled by the owner with less risk of running malware than a public workstation that can run any number of untrusted software components.

5.3 Discussion

The system requires users to have Internet access on their mobile device. One might argue that the user can then just as easily consume the services on the mobile device. However, some services are not well suited for use on a mobile device. For instance, entering lots of text or reading large documents is cumbersome on mobile devices. Moreover, the Internet traffic generated by the mobile device during authentication is very limited while some services might require a large bandwidth.

The system provides protection against phishing attacks. The smartphone retrieves the credentials from storage based on the URL of the service provider. If a slightly modified URL is used, the smartphone won't retrieve any credentials from storage. This prevents that the users credentials are submitted to an untrusted party.

Currently, the plugin is partly realized via a Java browser plugin. However, due to the many Java vulnerabilities being discovered, many browsers disable the Java plugin by default. However, with the pending introduction of the Web Cryptography API[6] and the widespread support of HTML5, the Java plugin could be replaced using technologies available by default in browsers. This would also ease adoption of other browsers.

The user's usernames and passwords should be securely backed up in case the smartphone is lost or stolen. Multiple strategies can be used. A user could maintain a local backup on his workstation, possibly encrypted with a master secret. A second strategy would be to have the smartphone application synchronize the encrypted usernames and passwords with an online backup sever. If the user holds multiple devices on which the application is installed, the devices can go through a pairing phase during which they exchange a secret key (e.g. similar to Firefox Sync). This secret key can, subsequently, be used to synchronize the passwords between the devices via an online (backup) server.

[6] More information can be found on `http://www.w3.org/TR/WebCryptoAPI/`

6 Conclusion

This paper presented a system for out-of-band password based authentication in a Web context. It is aimed at being compatible with the standard password based authentication system that is used by the large majority of Web based service providers to ensure user authentication. It increases the security by preventing malware running on the workstation from intercepting the username and password of the user without requiring any modifications to the authentication logic at the service provider. Users no longer need to remember their passwords which allows them to choose unique strong passwords. A prototype implementation is presented that has been tested and found compatible with several popular service providers, such as Facebook.

For future work, the presented approach can be made compatible with two-factor authentication systems commonly used by service providers. The user typically receives an SMS with a unique code that needs to be submitted together with the username and password. The mobile application can monitor the received SMS messages for the unique confirmation code that can, subsequently, be submitted to the service provider, together with the username and password.

Acknowledgements. This work was made possible through funding from the *MobCom* and *SecureApps* projects, granted by the Flemish *agency for Innovation by Science and Technology (IWT)*. The authors would also like to thank Kevin De Brucker for the valuable contribution of his master thesis.

References

1. Aloul, F., Zahidi, S., El-Hajj, W.: In: IEEE/ACS International Conference on Computer Systems and Applications, AICCSA 2009, pp. 641–644 (2009)
2. Adida, B.: In: Proceedings of the 14th ACM Conference on Computer and Communications Security, CCS 2007, pp. 48–57. ACM, New York (2007), http://doi.acm.org/10.1145/1315245.1315253, doi:10.1145/1315245.1315253
3. Lastpass (2013), https://lastpass.com/
4. Ross, B., Jackson, C., Miyake, N., Boneh, D., Mitchell, J.C.: In: Proceedings of the 14th Conference on USENIX Security Symposium, SSYM 2005, vol. 14, p. 2. USENIX Association, Berkeley (2005), http://dl.acm.org/citation.cfm?id=1251398.1251400
5. Pala, M., Sinclair, S., Smith, S.W.: PorKI: Portable PKI Credentials via Proxy Certificates. In: Camenisch, J., Lambrinoudakis, C. (eds.) EuroPKI 2010. LNCS, vol. 6711, pp. 1–16. Springer, Heidelberg (2011), http://dl.acm.org/citation.cfm?id=2035155.2035157
6. Boukayoua, F., Vossaert, J., De Decker, B., Naessens, V.: Using a Smartphone to Access Personalized Web Services on a Workstation. In: Camenisch, J., Crispo, B., Fischer-Hübner, S., Leenes, R., Russello, G. (eds.) Privacy and Identity Management for Life. IFIP AICT, vol. 375, pp. 144–156. Springer, Heidelberg (2012), http://dx.doi.org/10.1007/978-3-642-31668-5_11

7. Chiasson, S., van Oorschot, P.C., Biddle, R.: Graphical Password Authentication Using Cued Click Points. In: Biskup, J., López, J. (eds.) ESORICS 2007. LNCS, vol. 4734, pp. 359–374. Springer, Heidelberg (2007)
8. Wiedenbeck, S., Waters, J., Birget, J.-C., Brodskiy, A., Memon, N.: Int. J. Hum.-Comput. Stud. 63(1-2), 102–127 (2005)
9. Biddle, R., Chiasson, S., Van Oorschot, P.C.: ACM Comput. Surv. 44(4), 1–19 (2012), http://doi.acm.org/10.1145/2333112.2333114, doi:10.1145/2333112.2333114
10. Dolev, D., Yao, A.C.: On the security of public key protocols. Tech. rep., Stanford University, Stanford, CA, USA (1981)

Measurement Based Indoor Radio Channel Modeling and Development of a Fading Optimized Circular Polarized Patch Antenna for Smart Home Systems within the SRD Band at 868 MHz

S. Wunderlich, M. Welpot, and I. Gaspard

Abstract. Radio based smart home systems become more widespread over the next years. The reason for that is that the demands of home builders and homeowners are tending to green homes with high energy efficiency. Additionally smart homes deliver interoperability of home electronics and control panels in ways that enhance security, convenience, comfort and overall quality of life.

A drawback of radio based systems is the difficult predictability of signal quality, especially in indoor propagation scenarios due to the high variance in signal strength introduced by fading. So a priori knowledge of the expected path loss is crucial for an efficient dimensioning of the necessary radio network. This paper presents the results of an extensive measurement campaign to deduce realistic radio channel models for short range devices (SRD) applications at 868 MHz. Additionally, a concept to reduce fading effects is presented. This concept is based on the idea to mitigate fading caused by depolarization losses by using a circular polarized antenna in interaction with a linear polarized antenna. A proof of concept of this method was embedded in the previous mentioned measurement campaign.

1 Introduction

Indoor radio channels suffer usually from slow fading. That means that the characteristics of the channel vary only slowly in time compared with the duration of a data symbol. This can become a very dramatic effect in the communication between two (or more) radio devices. Especially if these devices are non-mobile devices, like in a smart home system, which keep a fix distance between each other and the receiver is located in a deep fade, since it will remain in this deep fade for a long (or unlimited) time interval [10]. A deep fade can significantly decrease the SNR of a communication link, which results in a reduced data rate or in a complete drop out of the radio link.

S. Wunderlich · M. Welpot · I. Gaspard
Hochschule Darmstadt, FB Elektrotechnik und Informationstechnik, Darmstadt, Germany
e-mail: ingo.gaspard@h-da.de

L. De Strycker (ed.), *ECUMICT 2014,*
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_16, © Springer International Publishing Switzerland 2014

Another effect which can also cause deep fades is fast fading. Fast fading occurs in indoor scenarios due to moving obstacles such as moving people. [4] In a fast fading channel, the transmitter may take advantage of the variations in the channel conditions using time diversity (which is often used by smart home systems). However, this paper does not cover the influences of moving people and objects on the radio channel, for further information on this topic, see e.g. [4].

The main reason for fading in indoor environments is multipath propagation. Transmitted radio waves are reflected and scattered on walls, ceilings, floors and other obstacles and can combine at the receiver in a destructive manner, due to different phase shifts of the arriving signal paths [5]. Another effect is depolarization. The scattered and reflected waves that contribute to multipath fading can also transfer energy from the transmitted polarization plane into the orthogonal polarization plane (called cross-polarization coupling) [1]. Such coupling occurs as a result of oblique reflections from the walls as well as due to scattering from indoor clutter, such as furniture [7]. This effect can cause significant degradation in signal quality if the transmitting and receiving antennas are using the same polarization. This effect contributes also to fading (polarization fading).

Since the delay spread of indoor channels is in the range of nanoseconds [3] and the bandwidth of the most indoor devices is relatively small (the SRD bandwidth @ 868 MHz is below 600 kHz), the channel can be assumed to show flat fading, so the coherence bandwidth is much larger than the signal bandwidth and the channel does not show a frequency-selective behavior.

These fading effects can lead to a significant reduced signal quality between sender and receiver and represent a significant problem for radio communication in indoor transmission scenarios.

One way to mitigate fading introduced by cross-polarization coupling is to use a circular polarized (CP) antenna in interaction with a linear polarized antenna. The benefit of using a circular polarized antenna at one end of the communication link is the irrelevance of the orientation angle of the polarized wave, so the CP antenna can receive radio waves at an arbitrary polarization plane and the effect of cross-polarization coupling can be annihilated. The only drawback of this method is the additional 3 dB polarization loss introduced by using a circular polarized antenna in combination with a linear polarized antenna. One example of the reduction of fading using a circular polarized antenna in comparison to a vertical linear polarized antenna can be seen in Fig. 1. So this method is based on the idea that all values are decreased by 3 dB but the mean value of the received power level is much higher than without this method. This is due to the fact that the influence of the fading spikes is drastically reduced.

Another positive effect of using a CP antenna is that the orientation of the linear polarized antenna does not influence the communication link anymore. This is particularly advantageous since the orientation of the antennas can seldom be guaranteed for the most indoor application scenarios. This approach is especially practical if a master-slave relationship between a central radio module and one or more radio subunits is given (like in centralized smart home systems). The central radio

module can use the CP antenna while the minor radio devices still use low-cost and space-saving linear polarized antennas.

To verify this theoretical concept a measurement campaign in six buildings was performed (see Section 2).

So the two main objectives of this paper are:

1. Basic design and implementation of a circular polarized patch antenna which should be used in an actual smart home system and which can be used for indoor radio channel modeling. This antenna is subject to further improvements regarding size and cost reduction. The antenna itself is a well known design (truncated-corner square patch), however the advantage of using a circular polarized antenna at one end of the link and a linear polarized antenna at the other end of the link in comparison to the conventional usage of linear polarized antennas at both ends should be proven on the basis of a comprehensive measurement campaign.
2. Development of a path loss prediction model which covers a versatile number of scenarios and can be used for a simple estimation of the expected path loss in an actual smart home installation. There does not exist many of these prediction models for this specific frequency domain which are based on an extensive measurement campaign.

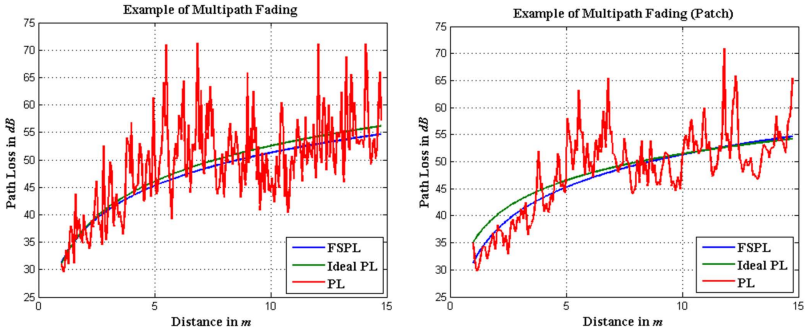


Fig. 1 a) A channel encounters strong fading using two vertical linear polarized antennas. The red graph (PL) represents the actual fading the channel encounters, the blue graph (FSPL) represents the theoretical free space path loss and the green graph (Ideal PL) represents the actual path loss the channel encounters without fading effects. b) Mitigation of Fading using a CP antenna in interaction with a vertically linear polarized antenna.

2 Measurement Setup and Procedure

2.1 Measurement Setup

The measurement setup is depicted in Fig. 2. The transmitting part of the setup is shown on the left side of the figure. The transmitting antenna is a cross polarized log-periodic antenna which is used to generate a horizontal and vertical polarized electromagnetic wave (denoted with XSLP9142 [8]). The circular polarized patch antenna can be used as an alternative transmitting antenna.

The two input ports of the cross polarized antenna are fed by two carrier wave (CW) generators at 873.8 MHz and 874 MHz at an output power of $P_{CW_{dBm}} = 14$ dBm. These slightly different frequencies were chosen to distinguish between the two polarizations states in the spectrum which were received by a CP patch antenna (or a dipole antenna as reference) and measured by a spectrum analyzer which stored it on a PC for further data processing.

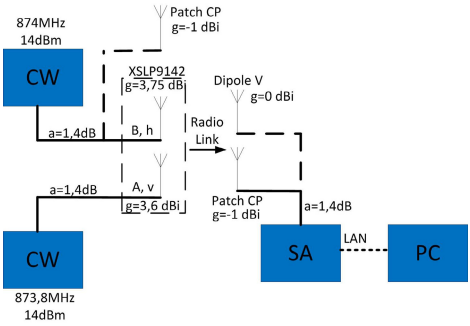


Fig. 2 Measurement and test setup for the indoor radio channel modeling and the circular polarized patch antenna

The height of the phase center of the transmitting and receiving antenna was $h = 1.16$ m. The transmitting antenna was located at a fixed position while the receiving antenna was moved along the measuring track during a single measurement run. A truncated-corner patch antenna design was used for the circular polarized antenna (see Section 2.2). The spatial resolution of a single measurement path was $d = 0.05$ m (distance between two measurement points) which results in a spatial sampling rate (SSR) of

$$\lambda = \frac{c_0}{f} = \frac{c_0}{874MHz} = 0.343m$$
$$SSR = \frac{\lambda}{d} = \frac{0.343m}{0.05m} = 6.86 \quad (1)$$

samples per wavelength which results in a sufficient detectability of fading spikes (where λ is the wavelength, c_0 is the speed of light and f is the operating frequency).

The path loss PL can be fairly simple measured by using the relationship:

$$PL = P_{Tx_{dBm}} - P_{Rx_{dBm}} \tag{2}$$

where $P_{Tx_{dBm}}$ is the transmitted and $P_{Rx_{dBm}}$ is the received power. Taking into account antenna gains and cable losses it follows:

$$\begin{aligned} PL &= (P_{CW_{dBm}} - a_{Tx} + g_{Tx}) - (P_{SA_{dBm}} + a_{Rx} - g_{Rx}) \\ PL &= P_{CW_{dBm}} - a_{Tx} - a_{Rx} + g_{Tx} + g_{Rx} - P_{SA_{dBm}} \end{aligned} \tag{3}$$

where $P_{SA_{dBm}}$ is the power at the spectrum analyzer and a is the cable loss and g is the respective antenna gain at Rx and Tx.

2.2 Circular Polarized Patch Antenna

To perform the measurements it was necessary to develop an antenna with the desired electrical parameters, with focus on an axial ratio < 3 dB (to achieve a good circularity of the antenna), accurate impedance ($Z = 50\Omega$) and resonance frequency ($f = 873.9$ MHz) matching.[1]

A truncated-corner square patch was chosen as basic design. The advantage of this kind of antenna is its simple geometry which is fast to simulate and easy to layout. Another advantage is that a circular polarized wave can fairly easy be created just by using a single-feed and without a 90° phase shifter. The impedance of the antenna can be matched by shifting the feed along one axis of the antenna.

Circular polarization can be achieved by combining two orthogonal modes with slightly different resonance frequencies. A single-feed patch with two opposite 45° truncated corners produces two orthogonal modes which resonate at different frequencies separated by almost 90° phase difference. The angle between the electric field vectors of both modes is given by the geometry of the antenna and is assumed to be 90° for rectangular patch antenna [2].

The antenna was designed using the simulation software "Sonnet Software". The simulation itself was performed in an iterative manner where the geometric parameters like patch size, corner truncation and feed point position were altered and simulated one after another until a satisfying solution was found. The resulting geometry of the antenna is depicted in Fig. 3a). The bottom side of the FR-4 material is completely metalized (copper). The resulting antenna is left-hand circular polarized (LHCP). However, since the other antenna in the communication link is always linear polarized, the handedness of the circular polarized field does not matter.

[1] The prototype antenna was designed for a resonance frequency at $f = 873.9$ MHz and not for the actual SRD frequency at $f = 868.3$ MHz.

Fig. 3b) shows the current density of the antenna at resonance frequency. The radial pattern shows the two orthogonal modes of the circular polarized wave with the highest current density in the center and the lowest at the corners.

Fig. 4 represents the $|S_{11}|$ parameter of the antenna. The two resonance frequencies represent the two orthogonal modes of the antenna. The antenna has a $|S_{11}|$ of 20.7 dB, a beamwidth of $\theta_{-3dB} \approx 60°$, an antenna gain of $g = -1$ dBi and a bandwidth of roughly $B = 13.2$ MHz. However, the bandwidth of the antenna is of inferior importance since the actual band of the SRD application is very narrow (600 kHz).

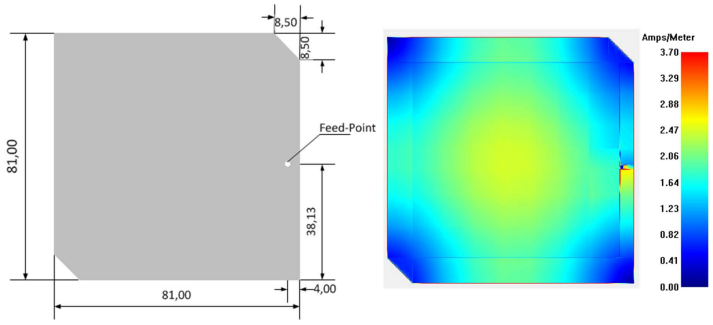


Fig. 3 a) Geometry of the circular polarized patch antenna (values in mm). b) Current density at resonance frequency.

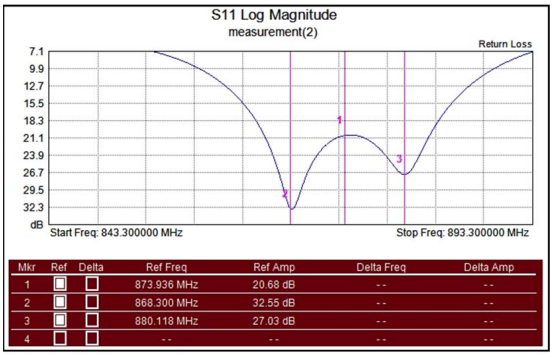


Fig. 4 $|S_{11}|$ of the realized CP antenna

2.3 Measurement Procedure

The measurements where performed in six buildings on the campus of the Hochschule Darmstadt. All of these buildings differ in parameters like material, geometry, usage etc. So a huge variety of different fading scenarios was achieved (see Table 1).

Table 1 List of buildings which were used during the measurement campaign

Building (name)	Type	Building materials	Designated usage
B11	One-story lightweight construction building	Wood	Lectures and seminars
B14	Two-storied shipping container architecture	Metal and wood	Lectures and seminars
C10	High-rise building (15 floors)	Reinforced concrete	Offices, lectures, seminars and labs
D11	Industrial building	Bricks and metal	Workshops, lectures and labs
D1617	Office building (5 floors)	Bricks and drywalls	Offices, seminars and labs

A section of a floor plan of one of the buildings can be seen in Fig. 5. The plan shows the entrance and the hallway of building B11. Basically two types of measurements were performed, LOS measurements within rooms and floors (denoted with yellow arrows) and NLOS measurements with walls, doors and furniture between transmitter and receiver (denoted with red arrows). The starting point and the direction of the single measurement runs were chosen randomly to guarantee an equal distribution of all possible propagation scenarios within the building.

3 Path Loss Prediction Model

The log-distance path loss model was used for the path loss prediction model. It is based on the assumption that the (mean) path loss PL is a function of distance d to the γ-th power

$$PL \sim d^{\gamma} \tag{4}$$

Where γ is the path loss exponent which indicates how fast path loss increases with distance. A logarithmic distance is used in the actual model, so the exponent γ becomes a factor and the formula reads:

$$PL = PL_0 + 10 \cdot \gamma \cdot \log_{10} \frac{d}{d_0} + X_g \tag{5}$$

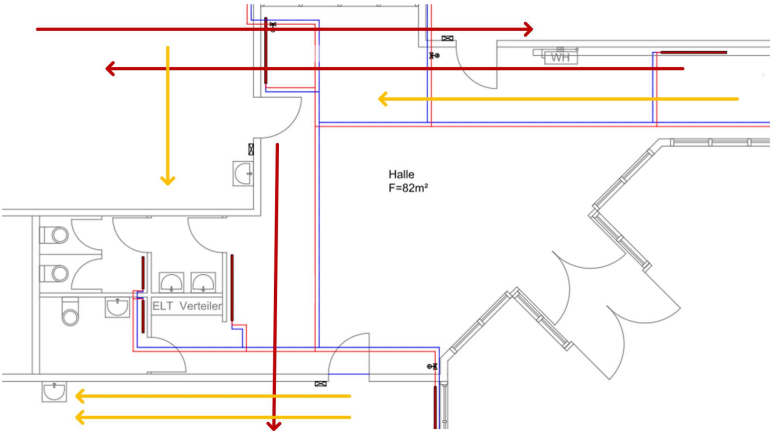


Fig. 5 Section of a floor plan of one of the buildings (B11) used for the measurement campaign

where PL_0 is the path loss at a certain distance d_0 and X_g is a normal distributed random variable which represents fading with zero mean and a standard deviation σ in dB. PL_0 and d_0 represent the path loss from the transmitter to the distance d_0 where the measurement starts [9]. d_0 can be freely chosen and was determined to be one meter for all measurements ($d_0 = 1$ m). The value of γ is dependent on the propagation scenario of the measurement, it is two ($\gamma = 2$) for free space path loss (FSPL), greater than two ($\gamma > 2$) when the path loss increases faster than in free space and smaller than two ($\gamma < 2$) when it increases slower than in free space.

For a homogenous environment γ can be assumed to remain constant for different distances between transmitter and receiver. However, if the propagation properties are changing (e.g. a wall separates transmitter and receiver) the path loss exponent γ is also changing. For this reason it is appropriate to divide the measuring track into several separate segments. γ can then be calculated separately for each of the segments. To compensate the additional path loss introduced by obstacles which influence the path loss exponent it is necessary to introduce new attenuation factors in Equation 5 [9]:

$$PL = PL_0 + 10 \cdot \gamma \cdot \log_{10} \frac{d}{d_0} + X_g + m \cdot FAF_{dB} + n \cdot WAF_{dB} \tag{6}$$

where FAF_{dB} represents a "Floor Attenuation Factor" and WAF_{dB} means "Wall Attenuation Factor". m and n are representing the number of floors respectively walls which are penetrated by the radio wave [9].

4 Results

The results of over 13,000 measuring points and 270,000 single measurements where recorded and processed to obtain a general fading model of the investigated buildings and an adequate data set to study the effects of the optimized circular polarized antenna in combination with a linear polarized antenna on fading.

Table 2 shows the cumulated results of the measurement campaign separated for every building. The first three columns are representing the path loss exponents for a vertically (V), horizontally (H) and circular polarized (CP) transmitted wave. The last six columns representing the standard deviation of fading for a LOS and NLOS connection between transmitter and receiver. All results are mean values of the results of the two available receiving antennas: a vertical polarized dipole and the circular polarized patch antenna.

Table 2 Path loss exponent γ and standard deviation σ in dB of fading for every building

Results	γ	γ	γ	σ	σ	σ	σ	σ	σ
Tx antenna	V	H	CP	V	H	CP	V	H	CP
Propagation				LOS	LOS	LOS	NLOS	NLOS	NLOS
B11	1.59	1.95	1.8	3.18	4.54	2.51	5.28	4.91	11.60
B14	1.94	1.70	1.71	4.59	4.11	4.58	5.45	5.19	7.25
C10	1.69	1.56	1.65	4.84	5.00	4.86	7.93	6.94	5.63
D11	1.99	2.07	2.03	3.86	4.28	3.3	5.50	5.50	4.15
D1617	1.95	2.3	2.17	4.3	5.02	3.67	5.33	5.12	6.01
Average Value	1.83	2.00	1.90	4.21	4.68	3.83	5.62	5.39	5.98

Some observations in Table 2 are remarkable:

- The path loss exponent is smallest if a vertically polarized wave is transmitted. One possible explanation for that is that the walls in a building can be treated as dielectric materials. Horizontal polarized waves can penetrate the walls if the angle of incidence fits to the permittivity of the wall material (Brewster angle phenomenon). Vertical waves do not suffer a similar effect and will be therefore reflected (according to [7]).
- The average path loss exponent is below two ($\gamma < 2$). This means that the path loss inside of the measured building increases slower than in free space. The reason for that is that the rooms and hallways of the buildings serve as a kind of wave guides due to reflections at floors and ceilings (see also the previously mentioned subitem).
- For a NLOS scenario the standard deviation is smallest if a horizontal wave is transmitted. A possible explanation for this phenomenon is the following: horizontal waves suffer less from reflections on walls (see the first subitem), so multipath fading becomes less important.

The results of Table 2 are independent from the used receiver antenna. Table 3 shows the standard deviation in dependency of the used receiver antenna (vertically polarized dipole or circular polarized patch antenna). Two points are remarkable:

- The standard deviation is the smallest for the transmitter/receiver antenna combination linear to circular (or vice versa). This outcome confirms the theoretical considerations introduced in Section 1.
- The combination circular/circular shows also good results, especially for LOS conditions. The explanation for this is that the energy of single-bounced reflected paths of the circular polarized transmitted wave disappears at the receiver since an odd number of reflections cause a reverse of the handedness of the circular polarized wave [6].

Table 3 Standard deviation σ in dB of fading in dependency of the used receiver antenna. Rx antenna: V = vertical polarized dipole, CP = circular polarized patch antenna.

Rx antenna	V	V	CP	CP	CP	V	V	CP	CP	CP
Tx antenna	V	H	V	H	CP	V	H	V	H	CP
Propagation	LOS	LOS	LOS	LOS	LOS	NLOS	NLOS	NLOS	NLOS	NLOS
σ /dB	4.90	5.40	3.52	3.96	3.83	6.00	5.67	5.52	5.12	5.98

A graphical representation of fading can be seen in Figure 6. It shows a percentile measure of the encountered fading. Most measurement values of the patch antenna are below the according measurement values of the dipole. For example the 90th percentile (see Figure 7) shows that 90% of the measurement values of the patch (LOS) are below a value of 5-6 dB, while 90% of the measurement values of the dipole are below a fading value of 7.5-8.5dB. The same situation shows for a NLOS connection, the fading of 90% of the measurement values of the patch are below 8-9dB, while 90% of the values of the dipole are below 9-9.5dB.

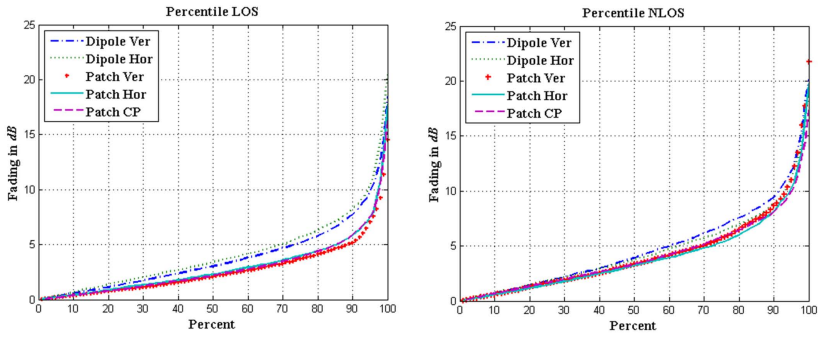


Fig. 6 Percentile measure of fading

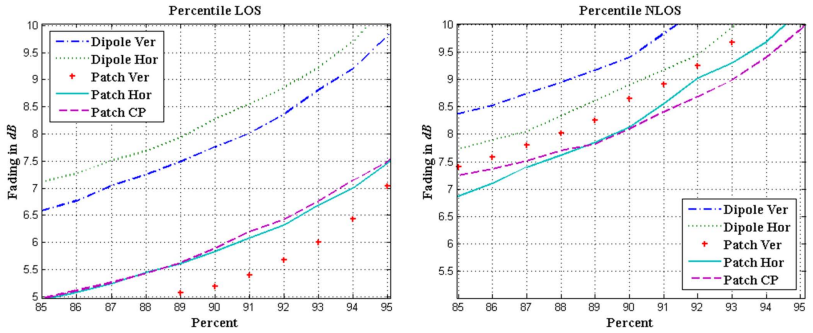


Fig. 7 Percentile measure of fading, segment of the 90th percentile

Two other papers which describe similar topics like this paper are e.g. [6] and [9]. [6] discusses different anti-multipath fading schemes using circular polarization under LOS conditions at 10 GHz. The results show that the mean amplitude fade level is reduced by $7-11$ dB by using CP transmission/reception. The results by using CP to linear transmission/reception are slightly less good.

[9] introduces a 914 MHz path loss prediction model with a similar approach like in this paper. Different propagation scenarios were considered (e.g. grocery store, retail store and two office buildings). The path loss exponent and standard deviation was determined within a value range of $\gamma = 1.81-5.04$ and $\sigma = 4.3-16.3$ dB, these relatively high values occur due to fact that the knowledge of the number of floors and walls between sender and receiver were not taken into account (see also Section 3). In relatively open environments (retail/grocery store) the path loss exponent is much smaller $\gamma = 1.81-2.18$ which is similar to the values which were encountered in this paper. The standard deviation of the most scenarios $\sigma = 4.3-8.7$ dB is also similar to the values presented in this paper.

5 Conclusion

The results of an extensive measurement campaign were presented. This measurement campaign was performed to create a measurement based fading model and to evaluate the possibility to fight fading using a circular polarized and a linear polarized antenna in an one-to-one transmission link within an indoor scenario.

The path loss exponent was measured within a range of values $\gamma = 1.83-2$ and the standard deviation with mean value of $\sigma = 4.24$ dB for LOS conditions and $\sigma = 5.67$ dB for NLOS conditions.

The method to combat fading using an optimized circular polarized truncated corner patch antenna has shown significant results. The results have shown that it is possible to mitigate fading by a notable amount, up to 3.5 dB for LOS and 1.5 dB for NLOS conditions (based on the 90th percentile), by replacing a linear polarized

antenna (which is normally used in indoor SRD communication systems) by a circular polarized antenna. This is of practical importance since this is a very cost-efficient, space-saving and easy-to-realize method.

This project (HA project no. 344/12-34) is funded in the framework of Hessen ModellProjekte, financed with funds of LOEWE Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz, Förderlinie 3: KMU-Verbundvorhaben (State Offensive for the Development of Scientific and Economic Excellence).

References

1. Cox, D., Murray, R., et al.: Cross-polarization coupling measured for 800 MHz radio transmission in and around houses and large buildings. Antennas and Propagation. IEEE Transactions on Antennas and Propagation 34(1) (1986), doi:10.1109/TAP.1986.1143714
2. Golio, M.: The RF and Microwave Handbook, pp. 6–127. CRC Press (2001)
3. Hashemi, H., Tholl, D.: Analysis of the RMS delay spread of indoor radio propagation channels. In: IEEE International Conference on Discovering a New World of Communications (1992), doi:10.1109/ICC.1992.268160
4. Horvat, G., Rimac-Drlje, S., et al.: Fade Depth Prediction Using Human Presence for Real Life WSN Deployment. Radioengineering 22(3), 758 (2013)
5. Molisch, A.: Wireless Communications, 2nd edn., 27 p. John Wiley & Sons Ltd. (2011)
6. Kajiwara, A.: Line-of-Sight Indoor Radio Communication Using Circular Polarized Waves. IEEE Transactions on Vehicular Technology 44(3) (1995)
7. Kyritsi, P., Cox, C.: Propagation characteristics of horizontally and vertically polarized electric fields in an indoor environment: simple model and results. In: IEEE VTS 54th Vehicular Technology Conference (2001)
8. Schwarzbeck Mess - Elektronik, XSLP 9142 Kreuzpolarisierte Breitband UHF-SHF Log.-Per. Messantenne Dual Polarized UHF-SHF Broadband Log.-Per. Test-Antenna 800 MHz... 3 (5) GHz (2013), `http://www.schwarzbeck.com/Datenblatt/manx9142.pdf`
9. Seidel, S., Rappaport, T.: 914 MHz path loss prediction for indoor wireless communication in multifloored buildings. IEEE Transaction on Antennas and Propagation 40(2) (1992)
10. Vireerackody, K.: Characteristics of a simulated fast fading indoor radio channel. In: Vehicular Technology Conference (1993), doi:10.1109/VETEC.1993.507051
11. Wong, K.: Compact and Broadband Microstrip Antennas, 162 p. Wiley Press (2002)

Influence of Bluetooth Low Energy on WIFI Communications and Vice Versa

Jeroen Wyffels, Jean-Pierre Goemaere, Bart Nauwelaers, and Lieven De Strycker

Abstract. In modern mobile devices such as smartphones and tablets, multiple wireless communication standards are supported. Since many of those standards operate at the licence free 2.4GHz ISM band, these standards use different modulation schemes and error correcting techniques in order to avoid interference when communicating over a wireless link (WIFI, Bluetooth, Bluetooth Low Energy...). For commercial use, these techniques are sufficient, but for industrial or medical use, in perhaps life threatening circumstances, all communications should perform at peak efficiency whilst not influencing each other. In this paper, a case study is presented where a healthcare setting is envisioned in which a WIFI network is deployed and used for medical purposes. We investigate the influence of a lot of transceiving Bluetooth Low Energy devices on the WIFI throughput, and vice versa. Conclusions and suggestions about the coexistance of both standards are also given.

Jeroen Wyffels · Jean-Pierre Goemaere · Bart Nauwelaers · Lieven De Strycker
KU Leuven Department of Electrical Engineering ESAT,
Kasteelpark Arenberg 10 B-3001 Heverlee - Belgium
e-mail: {jeroen.wyffels,lieven.destrycker}@kuleuven.be

L. De Strycker (ed.), *ECUMICT 2014,*
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_17, © Springer International Publishing Switzerland 2014

1 Introduction

In this paper, a discussion on the influence of WIFI communications on Bluetooth Low Energy communications and vice versa is presented. However the Bluetooth Low Energy standard (IEEE 802.15.1 [1]) has been specifically designed to operate in coexistence with all existing WIFI networks (IEEE 802.11x) and other 2.4GHz technologies, one cannot neglect the fact these standards all operate in the same frequency band, making communications in a possibly crowded frequency band less trustworthy as expected theoretically. This research can be of use in healthcare related situations, where possibly a lot of WIFI traffic is present (from, e.g., nurse call systems, Internet access for patients or care facility residents), but also a lot of Bluetooth Low Energy devices can be present. These Bluetooth Low Energy devices can be, e.g., incorporated in patient smartphones (commercial use), or wireless healthcare related devices as heart rate monitors, blood pressure monitors etc. which are emerging rapidly over the last months, making use of the Bluetooth Low Energy standard [2]. In Section 2, a general overview of the typical technical characteristics of the WIFI standard as well as the Bluetooth Low Energy standard is given, followed by Section 3 where the test setup of the research is presented. Section 4 elaborates on the results of the conducted interference tests between Bluetooth Low Energy and WIFI. In Section 5 a general conclusion and recommendation is presented.

2 Technical Characteristics

Bluetooth Low Energy makes use of the 2.4GHz Industrial Scientific Medical (ISM) band. There are 40 channels with 2 MHz channel spacing, of which three are dedicated advertising channels for device discovery, connection establishment and broadcast transmissions, whereas data channels are used for bidirectional communication between devices which have a connection with each other. These advertising channels have been assigned center frequencies that minimize overlapping with IEEE 802.11 channels 1, 6 and 11, which are commonly used in several countries because these channels are three Non-Overlapping DSSS channels (bandwidth becomes thus $\pm$22MHz). Bluetooth Low Energy uses an adaptive frequency hopping mechanism (AFH) on top of the data channels in order to face interference and wireless propagation issues, such as fading and multipath. This mechanism selects one of the 37 available data channels for communication during a given time interval. All physical channels use a Gaussian Frequency Shift Keying (GFSK) modulation. GFSK differs from FSK since the baseband pulses (-1, $+1$) pass as Gaussian filter first before going into the FSK modulator. Pulses are hence smoother and thus the needed spectral bandwidth can be limited. The used modulation index is in the range between 0.45 and 0.55. An index of 0.5 is close to a Gaussian Minimum Shift Keying (GMSK) scheme, so the radios power requirements are low. Two beneficial side effects of the used modulation index are the increased range and enhanced

robustness [3]. The physical layer data rate is 1 Mbps. The coverage range is typically several tens of meters.

IEEE802.11, however, uses Direct Sequence Spread Spectrum (DSSS) as modulation technique (more precise, IEEE802.11b uses DSSS, IEEE802.11g uses both DSSS and OFDM). The key benefit of this technique is the ability to share a single channel over multiple users, and should be resistant to intended or unintended jamming by making use of chipping codes. There are 13 WIFI channels available, with 5MHz spacing between them. For typical WIFI applications, different data throughputs are needed from a user point of view: high definition video streaming requires more bandwidth use compared to, e.g., sending an email [4]. For use in healthcare applications, and possible video streaming, a minimum bandwidth of 2Mbps must be maintained in all circumstances. In theory, it seems Bluetooth Low Energy and WIFI communications should be unaffected by each other in typical applications, due to the nature of the modulation techniques as well as the bandwidth specifications of both standards.

Table 1 Comparison between technical aspects of Bluetooth Low Energy and WIFI

Standard	Frequency [GHz]	Bandwidth [MHz]	Data rate [Mbps]	Modulation	Range Indoor	Range Outdoor
802.11b	2,412 - 2,472 [a]	20	5,5 / 11	DSSS	38	140
802.11g	2,412 - 2,472 [a]	20	6, 9 , 12, 18, 24, 36, 48, 54	OFDM or DSSS	38	140
802.15.1 BLE	2,402 - 2,480	2	1	GFSK	50	150

[a] Worldwide use except Japan

3 Materials and Methods

For this research, a WIFI router (LINKSYS WRT54G wireless G broadband router) is used for the WIFI communications. Since we want to test the influence of Bluetooth Low Energy devices, off the shelf Bluetooth Low Energy devices are used. For the test, we use 21 CC2540EM modules from Texas instruments [5][6][7]. According to the Bluetooth Low Energy standard, a Bluetooth Low Energy device can be found if it is in an advertising state. These advertising packets can be transmitted over 3 advertising channels out of which one is chosen randomly for each advertising packet. These advertising channels are channels 37 to 39 of the Bluetooth Low Energy Standard [1]. Since we simulate the presence of a lot of Bluetooth Low Energy devices present in a healthcare facility and since these devices typically don't send a lot of data packets per time frame when in a connected state, we let the CC2540EM modules broadcast advertising packets. This situation can for instance occur when a lot of visitors are present in a healthcare facility wearing a Bluetooth

Low Energy device in a connectable state. In this research, we fix the advertising channel to one of the three advertising channels. In this way, we force the advertisers to broadcast in the same channel and thus introduce 300% more traffic then normally present in this advertising channel, since all packets are transmitted by all advertisers in the same channel. The CC2540EM modules are programmed to broadcast an advertising packet each 20ms + 0-10ms, so on average, each module transmits $\pm$40 advertising packets per second.The test is performed twice: once with all Bluetooth Low Energy data traffic on channel 38@2426MHz and WIFI communications on channel 4@2427MHz, and a second time with all advertising data on channel 37@2402MHz and WIFI communications on WIFI channel 13@2472MHz. Since different WIFI channels are separated 5 MHz apart and Bluetooth Low Energy channels 2 MHz apart, choosing Bluetooth Low Energy channel 38 and WIFI channel 4 might result in two data streams interfering with each other since these channels are only separated 1MHz apart even though different modulation techniques are used, whilst when a choice is made to communicate over channels which are separated 70MHz apart, no interference is expected. In general, both tests are repeated 21 times, whilst each time the test is performed, an extra Bluetooth Low Energy device is activated, which also starts advertising in the same channel as the other devices. Each of the tests is performed long enough for a laptop connected with a 100Mbps port of the WIFI router to send a 575Mb file to another laptop (both HP Compaq 8710P) which is wirelessly connected to the same router. The WIFI router is not connected to the Internet and uses it's internal DHCP server to assign IP addresses to the laptops. By analysing the time it takes to send this 575Mb file and count the number of lost packets during this communication in function of the number of active Bluetooth Low Energy advertising devices, we get an idea of the influence of data traffic caused by the Bluetooth Low Energy devices on WIFI communications. For the WIFI analysis, the tool 'Wireshark'[8] has been used. For the analysis of the Bluetooth Low Energy advertising channel, a CC2540 module is programmed as a packet sniffer. This module is connected to a PC through a CC Debugger from Texas Instruments [9]. The PC is running the SmartRF Packet Sniffer software from Texas Instruments, by which an analysis of the number of received advertising packets can be made, as well as an indication about the number of packets with a Frame Check Sequence error (FCS). A Frame Check Sequence, in general, refers to the extra checksum added to a frame in a communications protocol for error detection. The most common known FCS technique is a Cyclic Redundancy Check (CRC). The FCS field contains a number that is calculated by the sender and is based on the data in the frame to be sent. The value of the FCS is added to the end of the frame-to-send. The receiver recalculates the FCS based on the received frame and compares this with the FCS number included in the frame. If the two numbers are different, an error is assumed.

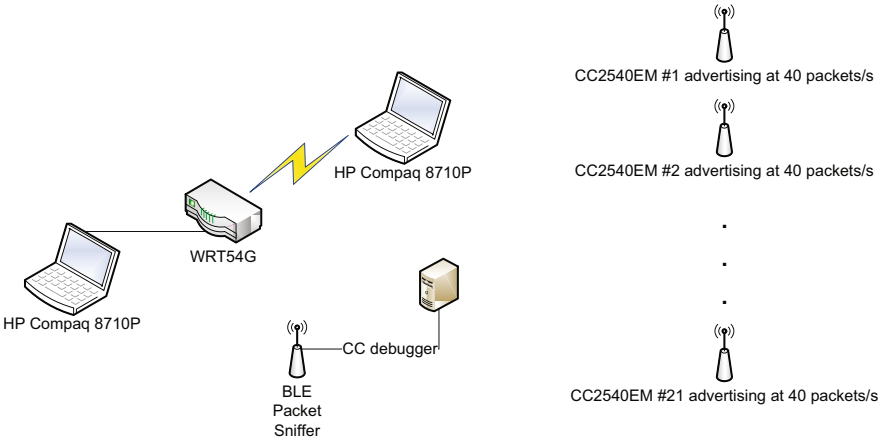


Fig. 1 Schematic overview of the test setup

4 Results and Discussion

4.1 Bluetooth Low Energy Advertising Channel 2402MHz, WIFI Communication Channel 2472MHz

In Table 2 an overview of the test results is given.

Table 2 Analysis of WIFI throughput in function of the number of active Bluetooth Low Energy Beacons (WIFI@2472MHz, BLE ADV@2402MHz)

Number of active BLE beacons	Number of WIFI packets sent	WIFI packets lost	throughput (Mbps)
1	642103	102	2,52
2	642203	105	2,58
3	642173	113	2,57
4	642499	97	2,50
5	642345	86	2,47
6	642306	103	2,58
7	642350	95	2,59
8	642260	105	2,62
9	642401	101	2,62
10	642187	97	2,62
11	642157	105	2,58
12	642514	76	2,45
13	642231	102	2,63
14	641929	88	2,62
15	641228	90	2,43
16	641306	103	2,66
17	641205	86	2,59
18	642106	100	2,54
19	642231	96	2,61
20	642321	98	2,63
21	641971	85	2,40

It seems that the influence of an increasing amount of Bluetooth Low Energy beacons on the WIFI communications can be neglected, which is what is expected since the communication channels are separated by 70MHz, and the bandwidth of the DSSS signal is 22MHz and of the advertising channel is 2MHz. In Fig. 2, a visualization of this experiment can be found:

- The WIFI throughput has not been influenced by an increasing amount of Bluetooth Low Energy beacons.
- The more Bluetooth Low Energy beacons become active, the less advertising packets per beacon per unit of time are received by the packet sniffer. This suggests that in this channel, more and more collisions between advertising packets from different beacons, occur. The amount of collisions, however, is not influenced by the WIFI communications but only by the number of active Bluetooth Low Energy beacons.
- The more Bluetooth Low Energy beacons become active, the more Frame Check Sequence errors occur. This complies with the fact more collisions are assumed in the advertising channel.
- The total number of received advertising packets in function of the number of active beacons per unit of time, drops from 40 packets per second to 22 packets per second. Also this leads supports the vision that a high amount of collisions occur in the advertising channel.

4.2 Bluetooth Low Energy Advertising Channel 2426MHz, WIFI Communication Channel 2427MHz

In Table 3 an overview of the test results is given.

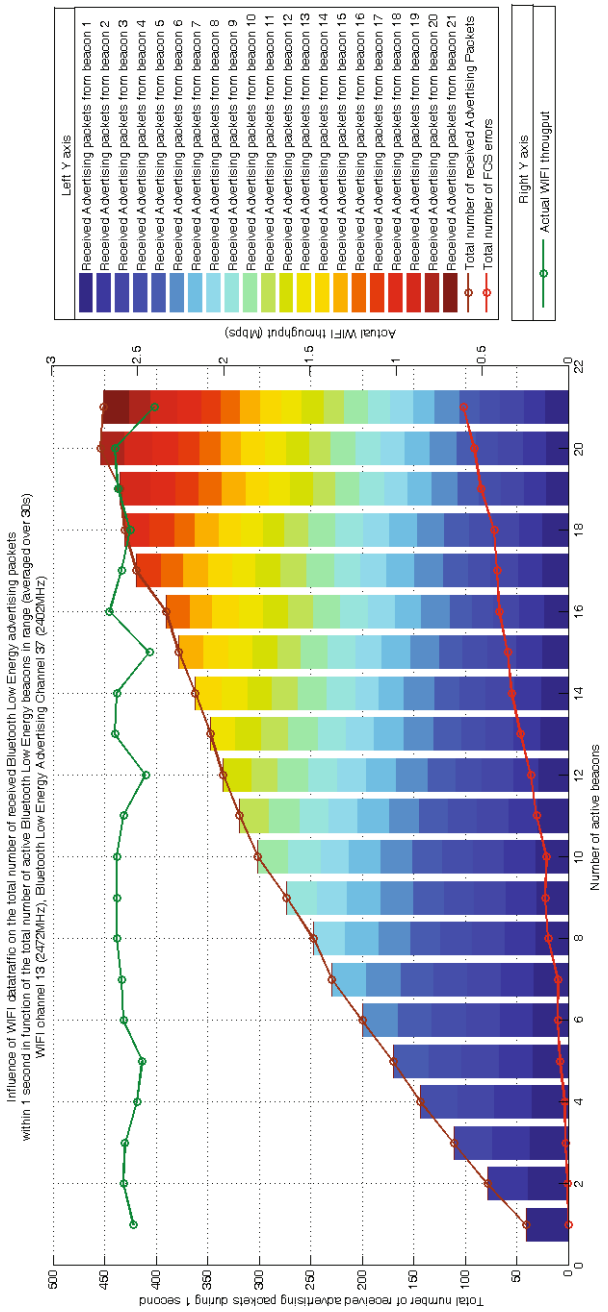


Fig. 2 Total number of received advertising packets within 1 second in function of the total number of active beacons in range. Bluetooth Low Energy advertising channel 2402MHz, WIFI communication channel 2472MHz.

Table 3 Analysis of WIFI throughput in function of the number of active Bluetooth Low Energy Beacons (WIFI@2427MHZ, BLE ADV@2426MHz)

Number of active BLE beacons	Number of WIFI packets sent	WIFI packets lost	Throughput (Mbps)
1	642861	106	2,50
2	642168	91	2,33
3	642105	97	2,26
4	642315	104	2,20
5	642311	100	2,19
6	642387	102	2,14
7	642189	96	2,11
8	642449	101	2,00
9	642225	92	1,94
10	642450	103	1,34
11	642857	132	1,83
12	642717	127	1,70
13	643700	209	1,62
14	642555	110	1,58
15	642750	143	1,58
16	643984	194	1,50
17	642381	125	1,46
18	642514	172	1,15
19	642521	97	1,45
20	643365	257	1,34
21	643982	140	1,30

It's clear the WIFI communication gets influenced by the presence of advertising Bluetooth Low Energy beacons. Since both channels are only separated by 1MHz, the bandwidth of the WIFI channel which is 5MHz and the Bluetooth Low Energy channel bandwidth is 2 MHz, both communications influence each other, however both systems don't fail completely. In Fig. 3, a visualization of this experiment can be found:

- The WIFI throughput is influenced by an increasing amount of Bluetooth Low Energy beacons. The actual data throughput drops by 50%.
- The more Bluetooth Low Energy beacons become active, the less advertising packets per beacon per unit of time are received by the packet sniffer. This suggests that in this channel, more and more collisions between advertising packets from different beacons, occur. The amount of collisions, however, is not influenced by the WIFI communications but only by the number of active Bluetooth Low Energy beacons.
- The more Bluetooth Low Energy beacons become active, the more Frame Check Sequence errors occur. This complies with the assumption that more collisions occur in the advertising channel.
- The total number of received advertising packets in function of the number of active beacons per unit of time, is much lower when a WIFI communication channel is in use nearby the advertising channel. This suggests that not only WIFI is influenced by the presence of Bluetooth Low Energy advertisers, but also vice versa: the presence of WIFI communication nearby an Bluetooth Low Energy advertising channel clearly limits the amount of advertising packets which actually reach the receiver.

The influence between Bluetooth Low Energy and WIFI is clear. This could be expected to some extent since the GFSK system and the DSSS system interpret each other as a source of noise (AWGN). The signal-to-noise ratio thus lowers, resulting in a higher bit error rate and thus a lower actual throughputs in both systems: the maximum theoretical data rate or channel capacity (C) in bps is a function of the channel bandwidth (B) channel in Hz and the signal-to-noise ratio (SNR): $C = B \cdot log_2(1 + SNR)$. This is the known Shannon-Hartley law, stipulating the maximum achievable data rate is directly proportional to the bandwidth and logarithmically proportional to the SNR.

5 Conclusions

In this paper, an overview of the possible influences of WIFI communications on Bluetooth Low Energy communications is presented. Depending on the field of application and the envisioned data throughput in the WIFI channels, it can be necessary to take precautions regarding channel assignments. Since Bluetooth Low Energy advertisers typically use three advertising channels randomly, it can be expected that only a small delay of device discovery might occur if WIFI channel

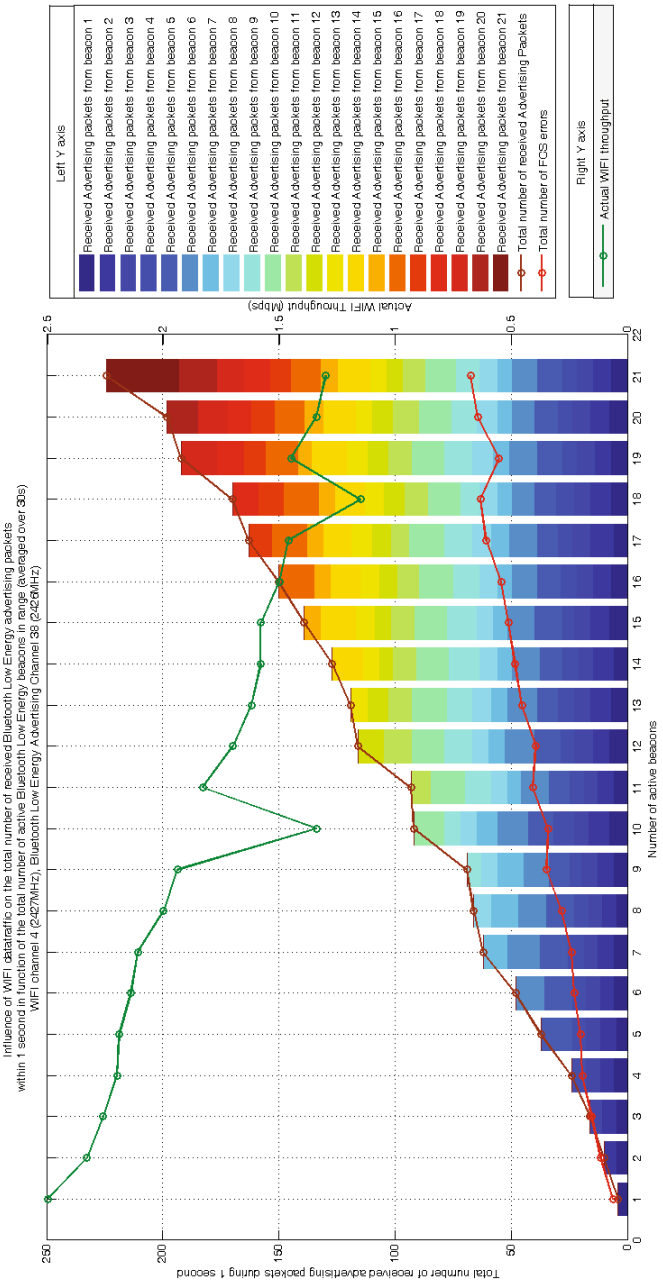


Fig. 3 Total number of received advertising packets within 1 second in function of the total number of active beacons in range. Bluetooth Low Energy advertising channel 2426MHz, WIFI communication channel 2427MHz.

4@2427MHz is used, since more or less half of the advertising packets in advertising channel 38@2426MHz might get lost, so in total roughly $5/6$ advertising packets are received by the receiver when searching for advertisers in all three available advertising channels. However, the decrease of the actual WIFI throughput can be more of an issue, again depending on the envisioned field of application. If only low data rates have to be achieved, one can randomly use whatever WIFI channel is available, but when high data rates need to be maintained for instance in streaming applications or commercial use, it is best to avoid WIFI channel 4 when a lot of Bluetooth Low Energy broadcasters are present. This choice assures the presence of Bluetooth Low Energy devices will not influence the actual WIFI throughput. As a final remark, one must be reminded of the fact this analysis was conducted with Bluetooth Low Energy advertisers fixed at one advertising channel. Since Bluetooth Low Energy uses AFH for data communication while in a linked state, no influence of Bluetooth Low Energy at WIFI or vice versa can be expected, keeping in mind that Bluetooth Low Energy has been developed for low data rate applications.

Acknowledgement. This Ph.D. research is funded by IWT, the Flemish agency for Innovation by Science and Technology (Baekeland mandate).

References

1. Specification of the Bluetooth System, Covered Core Package, Version: 4.0 (2010), https://www.bluetooth.org/en-us/specification/adopted-specifications
2. Wyffels, J., Goemaere, J.P., Verhoeve, P., Crombez, P., Nauwelaers, B., De Strycker, L.: A novel Indoor Localization System for Healthcare Environments. In: 2012 25th International Symposium on Computer-Based Medical Systems (CBMS), pp. 1–6 (2012), doi:10.1109/CBMS.2012.6266347
3. Couch, L.: Digital and Analog Communication Systems, 8th edn. Pearson (2013)
4. Cisco. Wireless lan design guide (2011), http://www.cisco.com/web/strategy/docs/education/cisco_wlan_design_guide.pdf (Checked on: November 29, 2013)
5. Texas Instruments, CC2540 Evaluation Module Kit. Texas Instruments, http://www.ti.com/tool/cc2540emk
6. Texas Instruments, Battery Board for System-on-Chips - SOCBB, http://www.ti.com/litv/pdf/swru241
7. Texas Instruments, CC2540 datasheet (2010), http://www.ti.com/lit/ds/symlink/cc2540.pdf
8. Lamping, U., Sharpe, R., Warnicke, E.: Wireshark User's Guide for Wireshark 1.11 (2004-2013), http://www.wireshark.org/download/docs/user-guide-a4.pdf (Checked on November 28, 2013)
9. Texas Instruments, CC Debugger User Guide. Texas Instruments (2010), http://www.ti.com/lit/pdf/swru197

Gestural Interfaces for Mobile and Ubiquitous Applications

Ionut-Alexandru Zaiti and Stefan-Gheorghe Pentiuc

Abstract. In the last years there has been a remarkable evolution in acquiring data on human gestures. Increasing efforts are focused on hand gestures as they provide an efficient interaction model for any application. This is more obvious in applications for desktops or gaming systems to be used indoors or relatively stand still activities in which the user does not leave the confines of a room or virtually defined space. However, mobile applications have not benefited the results of recent advances to the same extent. While most of the interaction we have with mobile devices is done using our hands, the model has remained unchanged. We suggest the use of natural hand gestures in both desktop and mobile applications and provide a starting point for achieving an engaging and realistic interaction while still retaining a connection with the surrounding physical space. We go over the work we have done towards this goal and provide some guidelines for designing applications based on hand gestures.

1 Introduction

In the age of ubiquitous computing [24] we are constantly surrounded by multiple computers and smart devices all presenting us with various choices of entertainment or other activities through different ways of integrated interaction. Laptops and desktop computers mainly use the touch pad, mouse and keyboard to allow us to provide input and commands, smartphones and tablets come with touchscreens. To these several other input methods are added as accessories or peripheral devices.

There are joysticks and other game consoles which are specialized to serve a particular purpose, such as those ever popular applications, video games [2].

Ionut-Alexandru Zaiti · Stefan-Gheorghe Pentiuc
University Stefan cel Mare of Suceava, 13, Universitatii, Suceava, 720229, Romania
e-mail: ionutzaiti@yahoo.com, pentiuc@eed.usv.ro

L. De Strycker (ed.), *ECUMICT 2014*,
Lecture Notes in Electrical Engineering 302,
DOI: 10.1007/978-3-319-05440-7_18, © Springer International Publishing Switzerland 2014

While these devices serve different purposes and seem to present various and very different ways of interaction they are all used with our hands. Most of the interaction we have with smart devices is done through our hands due to the impressive range of actions that can be performed from the simple moving of the mouse to the calibrated and synchronized text writing through a keyboard.

The newest entries to input devices, for gaming applications or otherwise, broke away from the relatively static model of previous ones. The Nintendo Wii[2] allowed the user to move around freely and interact with the application using motion gestures to simulate a real life action, such as in the game of tennis where the Wii sensor could be used as a tennis racket. This change contributed towards a more realistic and physically involving experience. The users were no longer bound to their seat and could explore a larger area while holding the game controller. Wireless communication allows the user to move around with a mouse as well but the interaction does not change based on the location of the user. In most cases a larger space is needed to perform various motions required to accomplish a goal.

Similarly, in the case of Microsoft Kinect[3], a vision based technique is used to acquire data on the whole body movement. What this means is that the user is presented with the opportunity to execute a wider range of gestures, with an increased accuracy. Both Wii and Kinect create a virtual game space in which the user can move relatively free.

We believe that given the interaction we have with our environment is mostly done through our hands, capable of a varied range of functions with different levels of precision [10], any gestural interface should be at least partially based on hand gestures. Furthermore, hands should not only be used as an actuator but as a sensor as well. We considered that a good opportunity to provide a complete experience in interacting with the virtual space is to use the objects around us as a method of input by reversing the process of hand pose and gesture recognition. In the case of mobile applications, even though most of the interaction is done through our hands, the gestures we are required to use are limited in range. We implemented several applications for smart mobile devices using data not only on the position of the finger on the screen but also on the flexion of the used fingers. Based on our work we provide a discussion and guidelines on using hand gestures in the context of both standstill and mobile applications.

2 Object Based Interaction

We argue that hand postures can inform on the object the user is manipulating. The result is transforming everyday objects into physical interfaces instead of using specialized equipment. While there are multiple choices for acquiring data on hand gestures [3, 7, 8, 18] we used a data glove which allowed a high degree of precision while the user could still execute gestures in any position he saw fit (which is not possible in vision based techniques, depending on the user's position

[2] Nintendo Wii (`http://www.wii.com/`)
[3] Microsoft Kinect (`http://www.xbox.com/en-GB/kinect`)

relative to the camera). We performed a study using the 5DT Data Glove[4] which includes 14 sensors: two for each finger (the knuckle and first joint) and four sensors that measure proximity between each pair of successive fingers. The data glove captures data at a frequency of 60 Hz. Users had to execute tasks of manipulation for 18 different objects, 6 basic shapes, each in 3 sizes, small, medium and large.

In this context, a gesture can be classified as either simple or complex. For simple gestures or postures the hand and fingers do not move and stand relatively still (such as in simply holding an object). Complex gestures require movement of the hand and fingers to any degree of complexity, such as in closing the fist (Figure 1).

Fig. 1 The gesture of closing the fist represented in four steps from left to right

Hand postures are represented as 14-dimensional vectors and they are classified using the Euclidean distance:

$$p = \{p_1, p_2, \ldots, p_{14}\} \in [0,1]^{14} \tag{1}$$

$$\|p - q\| = \left(\sum_{i=1}^{14} (p_i - q_i)^2 \right)^{\frac{1}{2}} \tag{2}$$

For the first task (object translation) the participants stood in front of a table and they were asked to pick up objects from their right side and move them to the left side. The order of the objects was randomly generated through a software application which instructed participants on their task. Hand posture data was captured during the action of picking up and moving the objects. For the second task (object exploration) each object had a series of digits from 1 to 6 inscribed on several locations of its surface. The participants were required to perform an exploration in order to identify a randomly generated sequence of the digits. We obtained recognition rates over 95% when discriminating between the given objects with data acquired from 13 participants [22]. Our results are in agreement with those reported by existing research [16].

To illustrate our results we applied them for video games, in part due to their popularity [2] and recent adoption of gesture input which allows us to reach a

[4] http://www.5dt.com/products/pdataglove14.html

Fig. 2 Using real objects to interact in a first person shooter: (a) a toy as a gun and (b) a cup as a grenade for a first person shooter [25]

large number of potential users. We also chose games because of their evolution, being one the most dynamic and adaptive type of application. As an example we give Counter Strike, a popular first person shooter, in which we used regular house hold objects to interact with the virtual world [25]. The user can pick up various objects from his surrounding and can use them in the game, thus creating a custom physical interface (Figure 2). We mention previous implementations of object based interfaces [9, 12, 20] where objects were specifically designed for a predetermined action. An advantage to this technique is the increased rate of recognition, each object having its own identity. However, as Sluis et al. [20] reported the users found the objects similar to a TV remote and separate from the environment. The objects had to be redesigned to appear more as decorative and create the impression of belonging to the environment.

2.1 The Holding Posture

To complement the analysis of our previous experiments we provide the results for a third task, the object holding posture, in which the participants were asked to identify the most comfortable holding posture for each object. The experimenter was present during the process and recorded the gesture data when the participant was ready. The task took approximately 5 minutes to perform and an average of 172 postures were gathered for each object, in order to make available the small variations which can appear during holding an object.

2.2 Recognition and Analysis

Following our analysis on the translation and object manipulation tasks we limited our use of classifiers to the nearest neighbour classifier, the k-nearest neighbour classifier and a multilayer perceptron (MLP) in order to recognize object size and

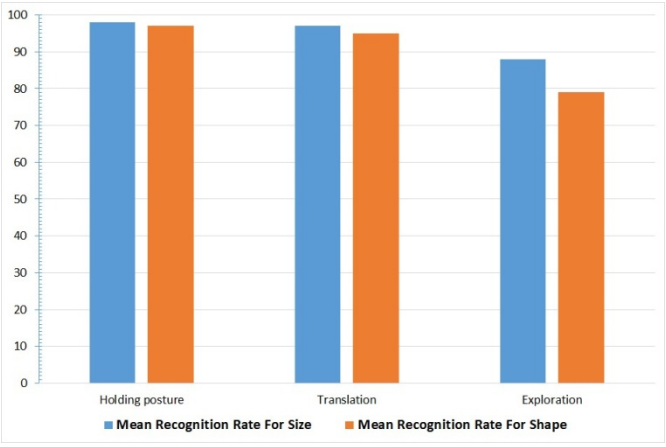


Fig. 3 Mean recognition rates for object shape and size using a nearest neighbour classifier applied on the raw data sets of all three object handling tasks

shape. Both nearest neighbour classifiers used a separating threshold of 0.1 which we determined to provide an acceptable compromise between data precision and recognition accuracy. The technique of calculating recognition rates for the holding posture was similar to the one for the previous tasks. For each object data set we randomly chose a fixed window of *w* postures which was used for testing while the rest of the data was used as the training set, a process which is repeated 100 times. The recognition rate is given by the formula:

$$\text{Recognition Rate} = \frac{\text{Correct Classifications}}{\text{100 Trials}} [100\%] \quad (3)$$

We note that we obtained similar and higher recognition rates for both the object size and shape using the nearest neighbour classifier and a window of only 10 postures (Figure 3), which is equivalent to the data collected in a sixth of a second, as opposed to the 30 posture window in our previous experiments. The recognition rates were of 98% for both the object size and shape. In comparison, in the case of the translation task the highest accuracy was obtained using the k-nearest neighbour classifier on the raw data (98% for both size and shape).

The higher accuracy in the case of the most comfortable holding posture can be explained partly through the specifics of the task, which ensure that each data set will provide a rather stable posture as opposed to a range of varied postures in the other tasks. Another factor is the low percentage of shared postures (Figure 4) compared to the other two tasks, 16% in the case of the holding posture as opposed to 22.3% in the translation task and 65.1% in the exploration task. The percentage of shared postures shows how many postures are common to two different object and it is calculated using the techniques established in our previous work [22].

We also wanted to determine how many times the most comfortable holding postures appeared naturally during the two other tasks, the object translation and

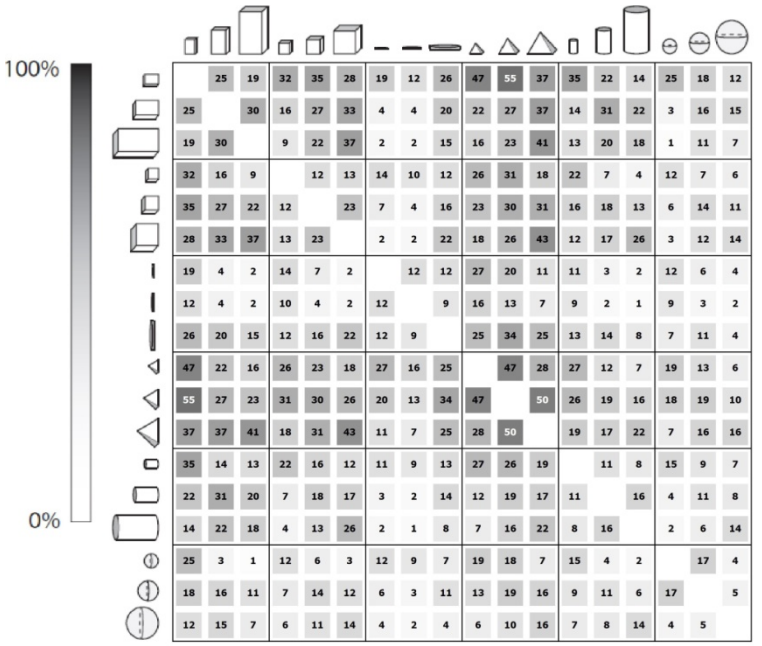


Fig. 4 Shared postures between the objects used in the experiment for the holding posture. A darker colour indicates a higher percentage shared postures between the two corresponding objects.

exploration. In order to calculate the percentage of holding postures used in the other tasks we averaged the holding posture data set for each object of each participant and we compared it to each posture of the corresponding data set from the other tasks. The process was executed for values of the separating threshold between 0.1 and 0.5 with an increasing step of 0.1. The results showed low percentages for the thresholds which provide a better resolution for the interaction postures of an object (Figure 5). Slightly higher percentages were found for the exploration task which provided a more varied range of postures as well as a larger data set, a total of 266595 postures compared to 57606 in the case of the translation task.

Given the size of the window, the used classifier (nearest neighbour) and the low percentage of shared postures between objects, real time recognition of the most comfortable holding posture for an object is possible and recommends it as a potential useful feature in object based interfaces. Considering the low percentage with which it is encountered in manipulating the objects and the high recognition rates, the holding posture offers some opportunities as an identification technique for both the object and the user.

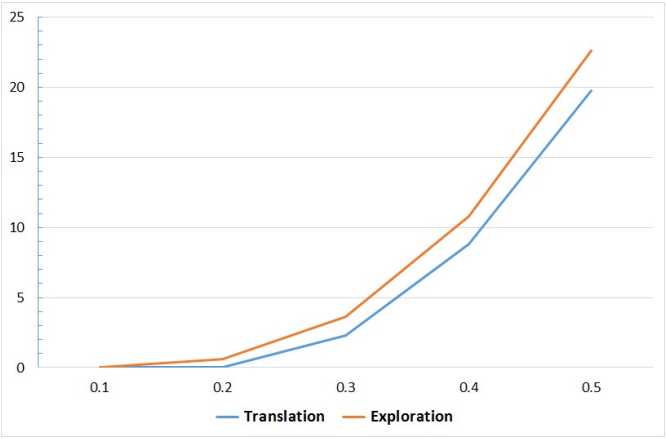


Fig. 5 The encounter percentage of the holding posture for each object in relation to the translation and exploration tasks for various values of the separation threshold

3 Gestural Interfaces for Mobile Applications

In recent years mobile applications have increased in complexity and in the count of their users. The touchscreen remains the main interaction tool for smart mobile devices due to the simple interaction model. We are allowed to directly manipulate displayed entities leading to one of the easiest to use interfaces [6]. However, touchscreens do come with some issues, such as selecting and manipulating a target on the screen, given that the resolution of the human finger is low. This problem is even more troublesome in the case of devices with a small screen estate. Most solutions to this problem originate from Fitts' law [4] or its variations [19]. Fitts' law gives us a measure of the time needed to select a target on a device for which we have two dependent parameters a and b as seen in equation 4. The time to select a target using a cursor is directly proportionate to the distance to the target (d) over the width of the target (w).

$$T = a + b \log \frac{2d}{w} \tag{4}$$

Various techniques such as offsetting the cursor [1, 13, 17, 23] have been developed to tackle the problem. Improvements have also been brought to the technology behind the touchscreen such as Tactus [21], allowing the touchscreen to become a deformable surface, or TapSense [5], detecting the part of the finger that was used to make contact with the touchscreen (the tip, nail or knuckle).

Even with such improvements the touchscreen still remains limited in the interaction model it allows, more so in relation to the rising complexity of mobile applications. We believe that using the data on hand gestures performed in the interaction with a smart mobile device would greatly benefit both the users and developed applications through the new dimension of supplied input. We implemented several applications using the simplest of hand posture data to show

the potential impact it could have. One of those applications was Paint where the user could draw given a set of simple colours. The addition we made was that the user could associate different colours to different fingers and also had various available actions (Figure 6). An example of such an action is identifying a colour on the screen for which the user could simply touch the screen with his little finger (the colour picker). Rubbing the screen with the side of the hand similar to brushing something off would produce clearing or erasing the screen in the touched area. Similar actions without the use of hand posture data would require either navigating a menu and selecting it or displaying the available actions on the screen and thus losing screen estate.

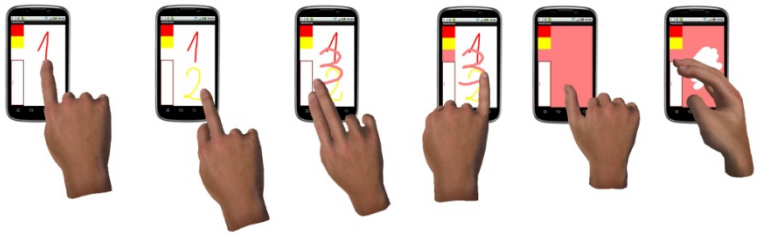


Fig. 6 Paint application using hand postures to differentiate between available actions

We previously mentioned a system in which objects became part of the interface for desktop gaming applications. Object based interfaces have been put forward before in multiple forms [9, 12, 20] but is there an argument for using them in mobile device applications. The main problem to appear with such a system is the handling of both the mobile device and the object at the same time. However, given the continuously evolving interaction between users and mobile devices [15] and the various uses of mobile devices such an interaction is possible and recommended. Current smart mobile devices have the ability to adapt to the user's requirements, being able to control large displays or smart TVs and even transform themselves into desktop systems. A precedent for mobile object based interfaces exists in applications such as Sphero[5] where the user can either control the Sphero ball with his mobile device as well as the other way around.

4 Creating Gestural Interfaces: Guidelines

Objects from an environment can act as a support for pre-existing gesture commands (as in the Counter-Strike application a cup could be used as a replacement for the grenade which had its own established gesture). However, human gestures are not only a tool for physically interacting with the environment. Supported by our work with object based and gestural interfaces we provide in Table 1 a summary of actions involving hands for such scenarios ranging from simple object manipulation to fine precision gestures.

[5] Sphero (`http://www.gosphero.com/`)

Table 1 Actions of the hand and fingers and application opportunities

Postures and gestures	Application opportunity
General object manipulation	The interaction we have with common objects all around us (e.g., drinking from a cup)
Precision gestures and complex actions	Using fine finger movements for actions that require precision (e.g., playing the guitar)
Culture-specific gestures	Waving, the 'ok' sign, thumbs-up, etc.
Communication oriented gestures	Conversational hand gestures [11] that assist inter human communication
Involuntary actions	Involuntary hand gestures and finger movements such as finger tapping

While the needed degree of precision and the variation of gestures depend on the application profile we consider that for a ubiquitous application to be complete its gestural interface must use gestures from all those listed in Table 1. The given categories are not to be viewed as independent of one another but overlapping in some contexts. Their goal is not only to cover the range of actions which can be executed by our hands but also to punctuate the varied dimensions of hand gestures and their essential role in our lives.

4.1 Object Manipulation Gestures

An example of such gestures is the Counter Strike application enhanced with hand gestures supported by objects found in the users' surrounding. The user has the option of customizing the interface himself by finding appropriately similar objects to those which are the subject of the given application. The shape and size of objects can be deduced from the gestures performed by the user while manipulating it [16, 22] which would mean that the objects required by the application should be recognizable and easily adjustable to by the user.

We performed a study on students aged 22-23 to provide some insight on how the correlation between physical and virtual environments is made, from the perspective of the available interactive objects. Gaming applications were chosen mainly due to their popularity [2]. We also took into consideration that the first person shooter and role playing game types provide an interaction involving an extensive range of virtual objects such as weapons, tools and household appliances. The participants were asked among others to provide examples of real life objects, with no restriction, that could be related to virtual objects in a game, to represent any entity. As it was expected most correlations were made based on both size and shape, such as using a round plate for a steering wheel or a blanket as a cloaking device in role playing games. However, there were some other results as well regarding perception of scale. Some of the examples did not directly correlate the size of the physical object with the one in the virtual game,

while still allowing a direct manipulation. One example was using a leaf as a boat, another to use a cabinet to represent a building, based on the similar shapes. This provides a potential direction of interaction with virtual entities by change of scale and using gestures defined for shapes as a command interface.

4.2 Precision Gestures and Complex Actions

These gestures refer to those actions which require a high level of accuracy or coordination, such as playing a musical instrument. While such activities may not always be directly required by users they are almost always implied through the requirements of the application itself. Some actions may not be replicated correctly based on the user's hand gestures without the accuracy of the input. However, even if required in order to create a more realistic and involving experience our recommendation is not to ground the application solely on them. From the first study we performed on object exploration and manipulation one of the secondary data we obtained was the reaction of the user. What we noticed was the level of fatigue in the case of finer gestures increased as the needed coordination in such cases requires a greater focus and thus puts a greater strain on the user.

While the first two categories regarded the technical side of gestures and recognition, the next three cover aspects of gestural interfaces that allow the user to feel as a part of the application, instead of just a separate element controlling it.

4.3 Culture Specific Gestures and Communication Oriented Gestures

Culture specific and communication oriented gestures introduce another dimension in gestural interfaces, that of symbolism. Most such gestures represent an idea or a state of mind, thus presenting the application with an input of information, which is much more than a command. The thumbs-up sign, which is executed by holding the fingers closed except for the thumb which is straight, can present multiple meanings depending on the context in which it is executed. It can give a confirmation to a direct question or it can present a state of being. Such gestures have an increased value to applications due to the added information they bring. We see a similar system used in current applications where, based on the current locale, we have present options in the user's language as well as other custom settings. The same concept can be applied to gestures based on the user's own culture, further increasing his integration in the application.

4.4 Involuntary Actions

Taking involuntary actions of users into consideration is a step even further from culture specific and communication gestures. These gestures no longer come from the user as a command, they are fully implied data or a raw display of information. Conversational hand gestures along with involuntary actions can be used to deduce the user's emotional state, such as nervousness or anxiousness. AffQuake [14] (based on the first person shooter Quake) or Relax to Win[6] are examples of pervasive games that use the emotional state of the player to influence various aspects of the game.

5 Conclusion and Future Work

In this paper we presented solutions we implemented for using hand gestures in applications for both desktop systems and smart mobile devices, gestures which were either standalone or supported by objects found in the surrounding environment. This allows users to build custom physical interfaces and have a natural interaction with applications. We complemented our previous study on object manipulation activities (translation, exploration) with an analysis of the holding posture for objects. Following our work and study on gestural interfaces we provided basic guidelines to promote using hand gestures as not only an input but also a provider of information and additional data such as emotional state. We believe our classification of hand gestures by their specific goal and particularities allows a better and a stronger ground for integrating gestures in the development of application interfaces.

Emerging technologies such as LeapMotion[7] or Myo[8] provide new ways of acquiring data on hand gestures which could bring benefits and create a momentum in gestural interface development. We believe that an interaction model based on hand gestures either standalone or supported by objects found in the user's environment can be the ground for truly ubiquitous applications. As future work we intend to extend our analysis of the object holding posture, specifically towards user independent recognition. We are also considering the implementation of an object based interaction model for mobile devices to determine the usability of such a model and the study of scale perception and interaction through shape oriented gestures in virtual environments.

Acknowledgements. This paper was supported by the project InteractEdu (Interactive gesture based system for the educational development of school-age children: applications in education, tourism and discovery of patrimony) 588/2012, co-funded by UEFISCDI, Romania and WBI, Belgium".

[6] Philips Design - Relax to Win (`http://www.design.philips.com/philips/shared/assets/design assets/pdf/nvbD/november2009/Getting emotional1.pdf`)
[7] LeapMotion (`https://www.leapmotion.com/`)
[8] Myo (`https://www.thalmic.com/en/myo/`)

References

1. Albinsson, P.-A., Zhai, S.: High precision touch screen interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2003, pp. 105–112. ACM, New York (2003), http://doi.acm.org/10.1145/642611.642631, doi:10.1145/642611.642631
2. Chatfield, T.: Fun Inc. Why Gaming will Dominate the Twenty-First Century. Pegasus Communications, Inc. (2011)
3. Erol, A., Bebis, G., Nicolescu, M., Boyle, R.D., Twombly, X.: Vision-based hand pose estimation: A review. Comput. Vis. Image Underst. 108(1-2), 52–73 (2007)
4. Fitts, P.M., Peterson, J.R.: Information capacity of discrete motor responses. Journal of Experimental Psychology 67(2), 103–112 (1964)
5. Harrison, C., Schwarz, J., Hudson, S.E.: TapSense: Enhancing Finger Interaction on Touch Surfaces. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST 2011), pp. 627–636. ACM, New York (2011), http://doi.acm.org/10.1145/2047196.2047279, doi:10.1145/2047196.2047279
6. Holzinger, A.: Finger instead of mouse: touch screens as a means of enhancing universal access. In: Carbonell, N., Stephanidis, C. (eds.) User Interfaces for All. LNCS, vol. 2615, pp. 387–397. Springer, Heidelberg (2003)
7. Huang, Y., Monekosso, D., Wang, H., Augusto, J.C.: A Concept Grounding Approach for Glove-Based Gesture Recognition. In: Proc. of the 7th Int. Conf. on Intelligent Environments (IE 2011), pp. 358–361. IEEE Computer Society, Washington, DC (2011)
8. Huang, Y., Monekosso, D., Wang, H., Augusto, J.C.: A Hybrid Method for Hand Gesture Recognition. In: Proceedings of the 8th International Conference on Intelligent Environments (IE), pp. 297–300 (2012), http://dx.doi.org/10.1109/IE.2012.30
9. Ishii, H., Ullmer, B.: Tangible bits: towards seamless interfaces between people, bits and atoms. In: Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI 1997), pp. 234–241. ACM, New York (1997)
10. Jones, L.A., Lederman, S.J.: Human Hand Function. Oxford University Press, Inc., New York (2006)
11. Krauss, R.M., Dushay, R.A., Chen, Y., Rauscher, F.: The communicative value of conversational hand gestures. Journal of Experimental Social Psychology 31, 533–552 (1995)
12. van Loenen, E.: On the role of Graspable Objects in the Ambient Intelligence Paradigm. In: Proceedings of the Smart Objects Conference, Grenoble, France, May 15-17 (2003)
13. Lu, H., Li, Y.: Gesture Avatar: A Technique for Operating Mobile User Interfaces Using Gestures. In: CHI 2011 Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems, pp. 207–216. ACM, New York (2011)
14. Magerkurth, C., Cheok, A.D., Mandryk, R.L., Nilsen, T.: Pervasive games: bringing computer entertainment back to the real world. Comput. Entertain. 3(3), 4 (2005), http://doi.acm.org/10.1145/1077246.1077257, doi:10.1145/1077246.1077257

15. Nova, N., Miyake, K., Chiu, W., Kwon, N.: Curious Rituals: Gestural Interaction in the Digital Everyday, Near future laboratory (2012), http://books.google.ro/books?id=XWbelwEACAAJ
16. Paulson, B., Cummings, D., Hammond, T.: Object Interaction Detection using Hand Posture Cues in an Office Setting. Int. Journal of Human-Computer Studies 69(1172), 19–29 (2011)
17. Potter, R.L., Weldon, L.J., Shneiderman, B.: Improving the accuracy of touch screens: an experimental evaluation of three strategies. In: O'Hare, J.J. (ed.) Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 1988), pp. 27–32. ACM, New York (1988), http://doi.acm.org/10.1145/57167.57171, doi:10.1145/57167.57171
18. Prodan, R.-C., Pentiuc, S.-G., Vatavu, R.-D.: An Efficient Solution for Hand Gesture Recognition from Video Sequence. Advances in Electrical and Computer Engineering 12(3), 85–88 (2012), doi:10.4316/AECE.2012.03013
19. Sears, A., Shneiderman, B.: High precision touchscreens: design strategies and comparisons with a mouse. Int. J. Man-Mach. Stud. 34(4), 593–613 (1991), http://dx.doi.org/10.1016/0020-7373, doi:10.1016/00207373(91)90037-8
20. van de Sluis, R., Eggen, J.H., Jansen, J., Kohar, H.: User Interface for an In-Home Environment. In: Hirose, M. (ed.) Human Computer Interaction, INTERACT 2001, Tokyo, pp. 383–393 (2001)
21. Tactus Technology, Inc. Taking touch screen interfaces into a new dimension, A tactus technology white paper (2012)
22. Vatavu, R.-D., Zaiti, I.-A.: Automatic recognition of object size and shape via Userdependent measurements of the grasping hand. International Journal of Human-Computer Studies (2013), doi:10.1016/j.ijhcs.2013.01.002
23. Vogel, D., Baudisch, P.: Shift: a technique for operating pen-based interfaces using touch. In: Proc. CHI 2007, pp. 657–666. ACM Press (2007)
24. Weiser, M., Brown, J.S.: The Coming Age of Calm Technology. In: Beyond Calculation, pp. 75–85. Springer, New York (1997), http://dx.doi.org/10.1007/978-1-4612-0685-9-6
25. Zaiți, I.-A., Pentiuc, Ș.-G.: Glove-Based Input for Reusing Everyday Objects as Interfaces in Smart Environments. In: Omatu, S., Neves, J., Rodriguez, J.M.C., Paz Santana, J.F., Gonzalez, S.R. (eds.) Distrib. Computing & Artificial Intelligence. AISC, vol. 217, pp. 537–544. Springer, Heidelberg (2013), http://dx.doi.org/10.1007/978-3-319-00551-5

Author Index

Alexandru, Nicolae Dumitru 1
Andrade, Maria T. 87
Andrei, Mihaela 11

Balan, Alexandra Ligia 1
Balan, Doru 143
Boukayoua, Faysal 23
Bruyneel, Karel 35

Caetano, Nídia 49

De Strycker, Lieven 109, 205
Dewitte, Karel 23
Dillenbourg, Pierre 153

Ferreira, Paulo 49

Gaspard, I. 193
Goemaere, Jean-Pierre 109, 205
Gorce, Jean-Marie 63
Graur, Adrian 135, 143
Guedes, Pedro 49

Harms, Hannes 49
Hutu, Florin 63

Janaszkiewicz, Anna 49
Joseph, Wout 73, 123
Juht, Toomas 49

Khoumeri, Aissa 63

Lapon, Jorn 181
Liu, Ning 73

Malheiro, Benedita 35, 49
Martens, Luc 73, 123

Naessens, Vincent 23, 181
Nauwelaers, Bart 205

Otebolaku, Abayomi M. 87
Ottoy, Geoffrey 109

Pakparvar, Mostafa 123
Pentiuc, Stefan-Gheorghe 217
Plets, David 73, 123
Potorac, Alin 143
Prelipceanu, Marius 135

Ribeiro, Cristina 49
Rimoldi, Bixio 153

Silva, António 49
Silva, Manuel 49

Tarabuta, Radu-Cezar 143
Tarniceriu, Adrian 153
Tarniceriu, Daniela 11
Trifina, Lucian 11

Uran, Christoph 167

Valauskaité, Jana 49
Van Den Berge, Sam 109
Villemaud, Guillaume 63
Vossaert, Jan 181

Welpot, M. 193
Wöllik, Helmut 167
Wunderlich, S. 193
Wyffels, Jeroen 205

Zaiti, Ionut-Alexandru 217