

Dynamic Adaptive Routing for a Heterogeneous Wireless Network

ERIC HSIAO-KUANG WU and YI-ZHAN HUANG

Department of Computer Science and Information Engineering, National Central University, Chung-Li, Taiwan

Abstract. This paper presents an integrated architecture of a Heterogeneous Wireless Network (HWN) and a dynamic adaptive routing protocol (DARP) for a HWN. To allow mobile users versatile communication with anyone or any device at any place and anytime, HWN integrates cellular network with an ad hoc network (independent Basic Service Set) in wireless local area network (WLAN) and reserves advantages of sizable coverage in a cellular network and high data rate in deployable ad hoc network. It also enlarges the scope of communication for ad hoc network and improves the throughput for cellular network. Consequently, nodes in HWN can communicate with each other or access Internet ubiquitously. We also address the routing issues for HWN, because the routing protocol for *HWN* is different from those used in cellular network or ad hoc network. The dynamic adaptive routing protocol establishes a better path for the source to arrive at the destination across multiple hops or cellular network and provides appropriate QoS (quality of service) in HWN.

Through simulation, we will demonstrate the merit of the HWN, proposed routing performance on HWN and overhead of control traffic. A performance analysis of the proposed protocol is depicted. The results of the analysis, and simulations, are presented and discussed.

Keywords: heterogeneous wireless network, wireless local network, cellular network, ad hoc network, multihop network, hybrid network, heterogeneous network, wireless network, routing, QoS, QoS routing

1. Introduction

1.1. Background, motivation, and challenge

Fueled by the new communication technology revolution and the information explosion, Wireless Personal Communication Systems (WPCS) [1] are now widespread all over the world. Mobile subscribers will demand to communicate with each other or access Internet services any time and anywhere. Therefore, wireless network services have to provide mobile users with ubiquitous communicating capability and information access.

A great number of wireless services have been developed for these needs and been classified by bit rate and service area. The satellite system delivers low bit rate service (typically 10 kb/s or less) over global coverage areas, while the Wireless Local Area Network (WLAN) [2] operates at Mb/s over hundreds of meters. The high-tier systems (cellular networks, like GSM [3]) have a coverage of may kilometers with a bite rate of around 10–20 kb/s, while the low-tier systems (cordless telephone systems [4], like CT2 and DECT) cover home or office areas with bit rate around 32–64 kb/s.

In most of the wireless networks, the WLAN and cellular network have offered human beings "wireless" information access and mobile phone services successfully. These two typical wireless access networks have their individual advantages and disadvantages. Current cellular networks provide transmission with low bandwidth, about 9.6 kbps (will be upgraded up to 100 kbps in GPRS and 384 kb/s in W-CDMA), even if they have large-scale coverage (about 1.5 km to 5 km). Thus, multimedia applications that require higher bandwidth cannot be well supported in current cellular network systems. Furthermore, cellular networks require fixed base stations to provide wireless access and communications must be relayed via these base stations. If there is no base station or the base stations fail, mobile users cannot stay connected with each other, even though the two subscriber nodes are near by. On the other hand, the ad hoc network in WLAN provides a higher bandwidth around 2 Mbps (will be upgraded to 11 Mb/s in 802.11b and 54 Mb/s in 802.11a), and better flexibility. In that way, ad hoc network can be deployed rapidly without infrastructure. It is suitable for mobile commuting, disaster recovery and multimedia transmission, but only covers a small-scale area.

Since these two wireless networks have different features and limitations, a single network cannot satisfy all users' requirements and needs. In order to fit the requirements of WPCS and support high utilization and popularization, it is essential to integrate these two different networks to have the options of a large-scale service area and high data transmission rate. Besides, to make good use of HWN, we propose a dynamic adaptive routing protocol (DARP) that contains a scheme supporting feasible QoS guarantee.

There are challenges in coordinating two network systems to operate on a global basis. Firstly, it is necessary to provide full compatibility of protocols and applications. The new amalgamated network should fit both existing users and existing operators. In other words, the existing applications should work on the new infrastructure properly and accurately. Secondly, the transparent networking should be provided. There will be numerous connections between distinct networks; therefore, functions including bridging or routing should be defined. Thirdly, the issue of effective user location management should be considered. The databases maintaining user location information in individual networks should exchange data interactively or be integrated into a single database to support global searches, pages, and accesses. Finally, for well-delivered multimedia applications, the integrated system should ensure the connection with quality of a service guarantee. The proper QoS control can promise the appropriate delay and bandwidth.

1.2. Related work

To address the insufficiency of a single network, researches were proposed on the internetworking between different networks. Proposed solutions include connecting LAN and wireless devices by Bridge-Routers (access point) [5], connecting cordless and cellular network for the 3rd generation [6], connecting Bluetooth and public networks (PSTN, Internet) [7], and a connecting single cellular network and a multihop network [10], etc.

In the paper [5], CASSIOPEE contains two major elements: the wireless communication provided by an impulse modulation of diffused infrared, and the backbone of a classical wired LAN. The Mobile Stations can communicate with each other through the backbone LAN via Bridge-Routers (Access Points). In [5], only a particular mobile station can communicate with other wireless devices within its radio scope. If the destinations are not within range or there is no BR point, the connections are invalid and cannot be established.

The paper [6] investigates a system that inter-networks between cordless and cellular systems, e.g., DECT and GSM. It states that the differences between DECT and GSM are correlative and complementary. Therefore, the integration can bring both users and operators significant benefits, and subscribers can utilize the same handsets at home and at the office as a cordless terminal within GSM coverage. The main assumption in the paper is that the whole system is of a cellbased infrastructure, so proper connections between base stations must be constructed.

Bluetooth Public Access (BLUEPAC) in the paper [7] integrates LAN and Bluetooth [8]. Users in wireless Bluetooth cells can access information in public areas, or even browse the Internet wirelessly. The paper also introduces a protocol concept of the "Bluetooth IP" which is based on Mobile IP and Cellular IP [9]. According to the ideas of "Bluetooth IP", BLUEPAC networks can support moving devices as well as static devices. The proposed infrastructure is also based on the structure of a cellular system. It requires nodes to connect to base stations before initiation information exchanges, and the radio devices have the limitation of short range.

The Multihop Cellular Network (MCN) proposed in the literature [10] combines the benefits of a conventional single-hop cellular network (SCN) and the flexibility of multihop network [11]. The analysis and simulation results show that the throughput of the MCN is higher than SCN and the throughput of MCN increases as the transmission range decreases. To demonstrate the feasibility of MCN architecture, a prototype based on the IEEE 802.11 Wireless LAN

[12] has been developed. A bridge protocol, which builds bridging tables between mobile stations and access points, is implemented for mobile stations to access the Internet properly through access points in multiple hops. In this paper, users can communicate with linking nodes via multihop way and with unlinking nodes via access points. However, access points are usually installed around a LAN environment with narrow coverage scope, so it might suffer from locating nearby access points.

To support essential, useful and powerful functionalities, it is required to have a routing protocol operating coordinately in the integrated network. The typical routing protocols of an ad hoc network can be classified into two main categories [13]: on-demand and table-driven. On-demand routing protocols are designed to reduce control overhead with minimal route acquisition latency. These protocols limit the bandwidth consumed by discovering routes to the destination only when a source has data to transmit. There are several examples for this approach, e.g., DSR [14], AODV [15], and TORA. Table-driven routing attempts to maintain updated information about the path among nodes in the network, thus the delay to find the destination node can be decreased. There are several protocols for this approach, e.g., DSDV, DBF, and so on. The ZRP (Zone Routing Protocol) is the hybrid protocol from on-demand and table-driven routing protocols. Hence, on-demand routing is source-initiated, but table-driven routing is destination-initiated.

1.3. Contribution of this paper

This paper proposes a HWN that integrates two typical network infrastructures and reserves its individual advantages. With the help of the dynamic adaptive routing protocol (DARP), there is no need to change hardware designs. Moreover, the proposed basic model of HWN may be extended or transferred to the new wireless technology or next generation mobile communications systems.

The dynamic adaptive routing algorithm addresses several issues:

- Connection establishment: rapid connection establishment, and functions of roaming between different networks;
- (2) Connection management: bandwidth requirement support, connection admission control, interconnection and traffic balancing between two networks, and compatibility for current networks and adaptation for prospective networks;
- (3) QoS support: bandwidth reservation to guarantee QoS for the real-time traffic connections.

DARP supports the above functions to accommodate applications without noticing the change of wireless infrastructure.

The paper is organized as follows. Section 2 specifies the architecture of HWM and describes the concepts of routing protocol. Section 3 defines detail routing functions, including

processes of register, routing, QoS, and maintenance. Section 4 presents mathematical analysis. In section 5, the simulation results are presented and analyzed. In the final section, we present conclusions and future work.

2. Heterogeneous Wireless Network (HWN)

2.1. HWN overview

The proposed architecture has two main components: one is Heterogeneous Wireless Network (HWN) and the other is dynamic adaptive routing protocol (DARP) over the HWN. We will describe the system topology, the architecture, the advantages, the concepts and the functions of routing for the environment of HWN.

HWN provides the solution that supports basic and advanced requirements to wireless personal communication systems. Generally, subscribers in HWN can initiate internet connection, voice conversation and interactive conference in an integrated manner. Consequently, HWN provides convenient and high-speed communications including global Internet, cellular network and wireless local area network.



F	igure	1.	HWN	topology



Figure 2. Protocol stack.

Technically, HWN integrates the features of cellular network and ad hoc network (figure 1). We assume that each node could be equipped with a cellular (e.g., GSM) interface and an ad hoc (e.g., wireless LAN) interface (figure 2). The basic idea of HWN is that mobile stations can communicate directly with each other or access cellular network through other mobile stations via multiple hops. It keeps the benefits of cellular network and incorporates the adaptability of ad hoc networks as well. Therefore, HWN supports all nodes in ad hoc network to connect to farther nodes via cellular network and nodes in cellular network can communicate with linking nodes with high-speed transmission service by ad hoc network without infrastructure. Therefore, HWN can make good use of infrastructure of cellular network and the feature of ad hoc network - fast and simple re-configuration. Besides, HWN can reduce the number of required base stations and improve the performance. Furthermore, it facilitates connections without base stations, multiple connections within the same cell, and high-speed packet transmission services. Last but not least, since base stations can help reduce the wireless hop count, paths are more stable and steady.

The detailed protocol stack is shown in figure 2. Each node could have both two interfaces, a cellular and an ad hoc, with different physical layers and MAC layers. Or it can have at least the ad hoc interface. The network layer is coordinated by DARP, composed by forwarding, route decision, user location management and registration. For supporting multimedia requirements, the bandwidth reservation, the queue management, and the other parts of QoS framework, are supported within the network layer. In the proposed stack, the MAC layer of the ad hoc network offers priority control for distinguishing real time traffic and data traffic. On the other hand, since the MAC layer of cellular network supports basic realtime (voice) traffic transmission, the QoS requirements (transfer delay and guaranteed bit rate) will be assured by the connection setup.

Nodes in the HWN have distinct behaviors from those in the cellular network. The dashed lines and solid lines in figure 3 demonstrate two distinct routing paths of the cellular network and HWN, respectively. If the source (S1) and the destination (D1) are located in the same cell, S1 can deliver



Figure 3. HWN architecture.

the packets through ad hoc network and then nodes between the source and the destination are used to relay packets to D1. This is multihop routing within a cell without the help of the base station. If the source (S2) and the destination (D2) are not located in the same cell but are neighboring nodes, S2 can connect to D2 by ad hoc mode directly without wasting the resource of the cellular network. If the source (S3) is in the cell coverage while the destination (D3) is not in any cell coverage, the traditional network will not support this type of connection. But in the HWN, packets can be sent to the base station first via a one multihop path, and forwarded to the base station where the destination resides, and afterwards packets will be forwarded to the destination, probably via another multihop path again. Based on this method, connections of different types can be supported by the HWN.

In order to support the HWN without severe changes or updates, a new routing protocol over the HWN is proposed to replace the traditional routing protocols. The new routing protocol must support the essential function first to find the way to the destination. Because the HWN is a hybrid network, the control packets in the proposed routing protocol (DARP) are designed to be able to pass through two networks in order to find the destination. Besides the fundamental function, the new routing protocol makes good use of the HWN to provide better routes. The routes are dynamically adaptive, which can be changed intelligently to fit different environments or conditions. Last but not least, the routing protocol can guarantee quality of service for supporting real-time traffic between two coordinated networks.

2.2. Dynamic adaptive routing protocol (DARP)

With the hybrid architecture of HWN, a new routing protocol is required to find a path in an intelligent way. Typical routing protocols in ad hoc are characterized into two types: one is on-demand routing, and the other is table-driven. The proposed routing protocol is based on on-demand routing because the advantages of on-demand routing fit HWN. Detailed reasons are described below.

- (1) Table-driven routing demands to record the whole users in the system to determine a route from the source to the destination rapidly. In HWN, there may be thousands or millions of online users. It would require a lot of work to maintain individual users' records and update periodically. Additionally, on-demand routing finds a route only when a transmission starts, so there is no need to record user data in the table.
- (2) Table-driven routing has lower control overhead, but the overhead grows as the connectivity increases or the percentage of transmission nodes decreases. While ondemand routing has higher control overhead, there are specific methods to reduce the overhead, such as cache.
- (3) The delay of on-demand routing is longer than that of table-driven routing, but it is still acceptable.
- (4) Bandwidth reservations in on-demand routing can be carried by existing routing packets thus the operations does

not require additional effort. While table-driven has to take other efforts to reserve bandwidth by sending packets before data transmission or exchange information of bandwidth in individual nodes.

DARP works in hybrid networks with different wireless physical characteristics. Routing on different networks requires nodes with two interfaces to forward packets on dissimilar physical mediums including control packets or data packets. Only nodes with WLAN devices can operate as routers to forward packets on wireless physical mediums. Two different devices that have different transmission range and diverse available bandwidth are adopted – one is cellular device and the other is WLAN device. A cellular device has larger radio range, but a WLAN device has more bandwidth available.

There will be many forwarding and rebroadcast activities, and thus a node may receive multiple copies of the same packet from various neighbors. When this occurs, the node drops the redundant packets and does not rebroadcast them. In addition, each routing path is restricted through cellular system only once: the first node on the path which has cellular device provides the path through the base station (BS). Later nodes cannot request the path through BS. Besides, all packets contain a field "hop count" which is incremented after passing through one node. If a node receives a packet that the number of the "hop count" is larger than the max hop count, the node will discard the packet without processing it any more. Due to the features of the broadcast, the node that broadcasts a packet can take the same packet rebroadcast as a passive acknowledge. Therefore, when the node receives a packet, it does not have to send a knowledge packet back.

DARP adopts the flooding approach for *Route Discovery* to find a route to the target node. The source node initiates a *Routing Discovery* packet as a single local broadcast packet and the intermediate nodes that receive this packet will rebroadcast it until the destination node receives it. As soon as the destination gets it, it will send a Routing Reply packet to the source and then a route is built.

The routing discovery process has 4 steps.

Step 1. The source node broadcasts a *Routing Discovery* packet. The *Routing Discovery* packet contains the following fields:

<source_addr, destination_addr, intermediate_set, cellular_pass, hopcount, sequence, ack_sequence, bandwidth, cellular_aid>.

The *intermediate_set* consists of a list pair of the *<de-vice_addr*, *device_type>*. Once a route crosses through multiple hops, the information of intermediate nodes' devices will be recorded. Whenever a packet is forwarded via cellular device, the *cellular_pass* bit will be set to indicate that the packet has passed through a cellular system. On the other hand, as the packet is broadcasted on WLAN device, it does not set the bit. Packets whose *cellular_pass* is set cannot be forwarded on a cellular device because it is not reasonable to



Figure 4. Routing formation (1).



Figure 5. Routing formation (2).

access cellular network more than once to find a specific destination. These packets can only be rebroadcast on the WLAN device as long as the hop count of this is lower than the max hop count (figures 4 and 5). Before a packet is forwarded or rebroadcast, the node's id (or IP address) is appended into the *intermediate_set* field.

Step 2. User location center will maintain individual subscribers' locations no matter whether they are located in cellular network coverage range or not. (It is specified in detail in section 3.2.) Whenever user location center receives the Routing Discovery packet, it will search the location table for the destination. If it finds a record with a BS responsible for the target destination, it will forward the packet to the BS. The BS will either send to the destination node or to the node that has assisted to register for the destination node for further forwarding. If it does not find any record containing the destination, it assumes that the destination is not in HWN system, and then forwards the packet to the gateway that connects internet or other systems. To reduce the overload of routing traffic, the BS does not broadcast packets to the nodes that could broadcast the packet on WLAN device for finding the destination further except the node which helps registration for the destination. Nodes that receive the Routing Discovery packet rebroadcast the packet on the WLAN device, and whether to rebroadcast the packet on the cellular device will rely on whether the *cellular_bit* is set on the premise that the hop count of the packet is smaller than the max hop count. And the node's node id (or IP address) is appended to the *intermediate_set* before the packet is forwarded.

Step 3. As the destination receives the first *Routing Discovery* packet, it will respond a Routing Reply packet that the route set is copied from the received *Routing Discovery* packet and the packet will be broadcast on the specified device. The Routing Reply packet contains the following fields:

<source_addr, destination_addr, intermediate_set, sequence, bandwidth, ack_sequence>.

Specifically, *destination_addr* is the original source address and *intermediate_set* is the reversion of the *intermediate_set* field in received *Routing Discovery* packet. The pairs in the field of *intermediate_set* provide information for intermediate nodes to specify the next node and the communicated device. The *ack_sequence* is set the same as sequence field in received *Routing Discovery* packet, and it is used for the original source to decide whether a received is the acknowledgement for certain *Routing Discovery* packet (figures 4 and 5). The duplicated *Routing Discovery* packets received later will be discarded and not coped. Then, the intermediate nodes will forward the packet according to the field of *"intermediate_set"* of the packet.

Step 4. As the source that has initiated the routing discovery process receives the *Routing Reply* packet, the Routing process succeeds. The source records the route set in the table and delivers data packets to the destination along the routing path by putting the routing path in the header of the data packet.

Detailed algorithm description is arranged in the following procedures. Moreover, the detail descriptions for the process of register, routing table and path maintenance, QoS support and adaptive adjustment are given below.

Routing Algorithm Procedure Procedure Routing

 if (has cellular device) set cellular_bit as 1
 send Routing Discovery Packet on available devices wait Routing Reply Packet
 <u>IDLE</u>
 if (Routing Reply Packet is received) routing success
 End Procedure

Procedure Receiving (Mobile Node)

if (hop count of the packet > max hop restriction) discard the packet

if (the packet received is the same as one of received packets before)

discard the packet

```
CASE of (received packet type)
Routing Discovery packet:
if ((packet received from cellular device) and (has not
enough bandwidth))
ł
      set cellular_aid as 1
      send packet on WLAN device
}
else if (has not enough bandwidth)
      discard the packet
else if ((has enough bandwidth) and (cellular_aid == 1)
and (has cellular device))
{
      set cellular_aid as 2
      append self IP in RouteSet
      send packet on WLAN device
}
else if (Destination == self IP)
{
      if (cellular_aid == 1)
             discard the packet
      else
   {
         write RouteSet (reverse from received packet)
         send Routing Reply packet along the RouteSet
   }
}
else if ((cellular_bit not set) and (has cellular device))
      append self IP in RouteSet
      set cellular_bit in the packet
      send packet on cellular device
      send packet on WLAN device
}
else
ł
  append self IP in RouteSet
  send packet on WLAN device
}
Routing Reply packet:
if (Destination == self IP)
{
   reserve required bandwidth
       routing success
   prepare data transmission
}
else if (self IP is not in RouteSet)
   discard the packet
else
{
  reserve required bandwidth
       send packet to next hop in RouteSet on
       corresponding device
}
Data packet:
if (Destination == self IP)
```

```
{
      release reserved bandwidth
      data transmission success
    else if (self IP is not in RouteSet)
       discard the packet
    else
    ł
      store the data packet
          prepare to send the packet to the next hop in
          RouteSet on corresponding device (on wireless -
          send RTS/CTS before data; on cellular - send in
          proper slots) release reserved bandwidth
    End Case
End Procedure
Procedure Receiving (User Location Center)
    CASE of (received packet type)
    Routing Discovery packet:
        send packet to the helping node
    Routing Reply packet:
        send packet to the next hop in RouteSet
```

send packet to the next hop in RouteSet

The Register process is operated as the system initiates or

Max_Register_Timer expires. Since the routing protocol finds

not only nodes in cellular network but also nodes in ad hoc

network, the nodes in ad hoc network are required to register

to attach a BS periodically. Other nodes in Internet or other

cells can find a node within certain BS according to the regis-

tered record in user location center. The Register process has

Step 1. The source node initiates the register process and

checks the existence of cellular device. If the node has active

cellular device, it registers to BS directly. Else it broadcasts

Register Discovery packet via the WLAN interface. The Reg-

<source_addr, intermediate_set, hop count, sequence>.

registering for it, there is no destination_addr field specified.

If a node without cellular device receives the Register Discov-

ery packet, it appends its address in the intermediate_set and

rebroadcasts Register Discovery packet using WLAN device

Step 2. As a node that has cellular device and is ac-

tive within cellular network receives the Register Discovery

Since the source does not know which node can assist in

ister Discovery packet contains the following fields:

Data packet:

End Case

3. Detail functions

End Procedure

3.1. Register

4 steps.

(figure 6).

224



Figure 6. Register process (1).



Figure 7. Register process (2).

packet, it will respond a *Register Discovery Reply* packet back to the source. The *Register Discovery Reply* packet contains the following fields:

<source_addr, destination_addr, intermediate_set, sequence, ack_sequence>.

The *ack_sequence* is set the same as the sequence field in received *Register Discovery* packet, and will be used for the original source to decide whether a received packet is the acknowledgement for certain Register Discovery packet (figure 6).

Step 3. As the source receives a first *Register Discovery Reply* packet, it responds a *Register* packet to the destination in the *Register Discovery Reply* packet, because the first received packet has the shortest delay time. The Register packet contains the following fields:

<source_addr, destination_addr, intermediate_set, sequence>.

The intermediate nodes will be described by the track recorded in the packet. Besides, if the source receives duplicate Register Discovery Reply packets later, it will discard packets (figure 7). Step 4. As the destination receives the *Register* packet, it will assist the source to register through the BS and return *Register Reply* packet back to the source. Once the source receives *Register Reply* packet successfully, the Register Process completes (figure 7). If the source does not receive reply for a specific period after sending Register packet, this indicates that the current Register Process fails, and the source could commence another Register Process immediately or wait MAX_Register_Timer to expire. Whenever Max_Register_Timer expires, nodes must start again the Register Process to update information on user location center.

3.2. Routing table maintenance

In DARP, each node maintains two types of routing tables. One table stores header information for packets that have been sent, and another table stores information for on-going connections. The first table contains the following information:

<source_addr, destination_addr, sequence, timeout>.

The table is used for waiting acknowledgements or avoiding sending the same packets twice. When the source has sent *Routing Discovery* packet or *Register Discovery* packet, it has recorded necessary information in this table and set a timeout. Within certain periods, if the source receives an acknowledgement for a record in the table, the acknowledgement is valid for the discovery. As the time expires, the record will be deleted, and thus later received acknowledgements are invalid. When the intermediate node forwards packets, it also records necessary information in this table. Within certain period, as received packets contain the same sequence number as before, the packets are discarded without the need to rebroadcast. The second table contains the following information:

<source_addr, destination_addr, sequence, previous hop, next hop, bandwidth>.

The table is used for managing resources and maintaining the routing paths. When nodes receive a *Routing Reply* packet, they would record the connection information in the table. Later when the connection time expires, the corresponding record is deleted. Due to no packets received from previous-hop node or next-hop node within at a certain period, the node will consider that the link to previous-hop node or next-hop node is broken due to movement or shutdown or else. The link broken detection determines to initiate the process of routing path maintenance described later.

Additionally, user location center, like VLR or Gateway, maintains a registration table that manages locations and information for individual subscribers. When user location center receives a *Register* packet, it will record the destination node's IP or MAC address and information, such as Phone number. When received packets are required to send to a certain destination, the packets are sent to the nodes registering for the destination first. If there is no record about the destination, the packets are sent to Internet or discarded.



Figure 8. Path maintenance (1-1).



Figure 9. Path maintenance (1-2).

3.3. Routing path maintenance

Movements of source or destination or intermediate nodes may break an active connection in HWN. It is necessary to perform a recovery method to re-build a path to the destination for continuous transmission.

The routing path maintenance process has 3 steps.

Step 1. As an intermediate node detects that the link to previous-hop node in this connection remains and the link to next-hop node fails, then the node broadcasts a *Routing Discovery* packet with minor modification to rediscover the destination. The packet contains the address of the original source and the address of the node that sent the packet (figures 8 and 10).

Step 2. Neighbor nodes rebroadcast this packet via the operation of HWN Routing. As the destination receives the packet, it will broadcast a *Routing Reply* packet (figures 8 and 10).

Step 3. If the node that has sent *Routing Discovery* packet receives the reply packet, it changes the *intermediate_set* by combining new path and old path information and packets later will pass through new path (figure 9). If the original source also receives the reply packet, this means that the destination is closer to the original source, and then the original source will change the connection information in the table by



Figure 10. Path maintenance (2-1).



Figure 11. Path maintenance (2–2).

replacing old *intermediate_set* with new *intermediate_set* and packets sent later will pass through a new path (figure 11).

3.4. QoS support

For real-time traffic, such as voice or video transmission, a scheme over basic routing is designed to guarantee quality of service for subscriber's delay and bandwidth requirements. Three key properties have been introduced to support the QoS.

(1) Bandwidth reservation. Each node which receives Routing Discovery packets will check whether it could serve the bandwidth required for the connection. If the node can serve the requirement, the packets will be forwarded further; otherwise, the packets are discarded. As long as the corresponding Routing Discovery Reply packets of real-time connection are sent back to the source from the destination, they can reserve required bandwidth on nodes along the path, and thus the relatively available bandwidth in the nodes is reduced. And these nodes record the information of real-time connections containing the amount of bandwidth reserved and the duration time of the connections. As the duration time is up, the connection is closed and the reserved bandwidth will be released and available again.

If the path passes across cellular network, there are some conditions in routing that need to be noted as below.



Figure 12. Routing with QoS support.

- (a) If the packet is received from BS and the bandwidth required for the route path can be served at this node, and cellular_aid is not set before, the node appends its information into the Routing Discovery packet and the packet is rebroadcast on the WLAN device to continue the process (figure 12).
- (b) If the packet is received from BS and the bandwidth required for the route path can not be served at this node, and *cellular_aid* is not set before, it is because the node which helped to register can not supply the quality service required. Then it sets the *cellular_aid* as 1 that indicates the route path needs another node to access cellular network. The node does not need to append its information into the Routing Discovery packet and the packet is rebroadcast on the WLAN device to continue the process (figure 12).
- (c) And if the received packet contains *cellular_aid* set as 1 and the node is active in cellular network and has enough bandwidth, it will indicate that the node can help for the requirement of quality of service. The node appends its information into the *Routing Discovery* packet, set the *cellular_aid* as 2 and the packet is rebroadcast on the WLAN device.

To operate these steps, the reservation path of *Routing Reply* packet will go through cellular network by the node that sets the *cellular_aid* as 2.

(2) *Priority*. To decrease queuing delay, real-time traffic related packets have higher priority than packets of data traffic. Thus, the routing packets and content packets for real-time traffic have less waiting time in queues. The other way to raise priority is that the packets related with real-time traffic have shorter back-off time for increasing the probability of early wireless medium access.

(3) *Queue reservation*. To guarantee queuing delay more accurately and avoid that the real-time traffic is influenced by other data traffic, we place a limitation on queues. If the transmission time of normal data exceeds the coming starting time of real-time data, the transmission of normal data suspends. Actually, there is another method to enforce the higher priority for real-time traffic.

Table 1 The parameters for mathematical analysis.

Attribute	Description
N	Number of nodes
τ_1	Probability that a station (real-time) transmits in a generic slot time
τ_2	Probability that a station (non-real-time) transmits in a generic slot time
P_1	Probability that a transmitted packet collides (real-time)
P_2	Probability that a transmitted packet collides (non-real-time)
$P_{\rm tr}$	Probability that time there is at least one transmission in a slot
P _{a1}	Probability that a specific station (real-time) transmit successfully
P_{a2}	Probability that a specific station (non-real-time) transmit successfully
Ps	Probability that a transmission is successful
hc	Hop count

3.5. Adaptive adjustment

Due to differences between a cellular network and an ad hoc network, several adaptive adjustments can be applied for different conditions.

(1) To satisfy individual distinct requirements or maintain the system efficiency, the routing protocol can allocate different paths for adaptive adjustment. For example, if a node requires a transmission of short delay and low bandwidth, the routing protocol can choose the cellular way for the node. On the other hand, if a node requires a transmission of high bandwidth, such as multimedia or multicast or broadcast transmission, the routing protocol can choose the multi-hop way for the node.

(2) If the network administrator of HWN desires to accommodate more simultaneous connections in cellular network, it can increase the max hop count to allow more connections using wireless multihop way but not cellular way.

(3) For better load balancing, HWN can confine a connection number to balance traffic between a cellular network and an ad hoc network. This can help avoid the congestion occurred in one network while another network is unoccupied.

(4) As there are few users in HWN, the max hop limit can be increased to raise the attaching probability. That is because the lower density makes higher hops, and to access the cellular network requires going through more nodes. However, the larger max hop count could waste more bandwidth in an ad hoc network due to many redundant re-broadcasting packets.

4. Mathematical analysis

Our objectie is to propose a better network architecture to provide ubiquitous access and a more effective QoS routing for real time traffic. Therefore, the analysis will address two topics: (1) the probability of successful real-time packet transmission and non-real-time packet transmission, (2) the end-to-end packet delay in HWN. The parameters used in this section are listed in table 1.

4.1. Successful transmission analysis

We describe the analysis in [16] and [17] briefly at first; consider a fixed number N of contending stations and let τ be the probability that a station transmits in a generic backoff slot time. The value of τ can be calculated via equations as follows:

$$\tau = \frac{2(1-2p)}{(1-2p)(W+1) + pW(1-(2p)^m)},$$
 (1)

$$p = 1 - (1 - \tau)^{N-1},$$
(2)

where p is the probability that at least one of the N-1 remaining stations transmits. W denotes the initial backoff window, and m specifies the maximum backoff stage. Given n active stations and the probability P_{tr} that there is at least one transmission in a slot time t. P_s denotes the probability of a successful transmission, and P_a indicates that a specific station transmit successfully.

$$P_{\rm tr} = 1 - (1 - \tau)^N, \tag{3}$$

$$P_{\rm s} = \frac{N\tau (1-\tau)^{N-1}}{P_{\rm tr}},$$
 (4)

$$P_{a} = \frac{\tau (1 - \tau)^{N-1}}{P_{tr}}.$$
 (5)

Since real-time packets have higher priority than non-realtime packets in HWN, there are two individual initial backoff windows and maximum backoff stages. Based on the equations described above, we derive the extended equations as follows:

$$\tau_1 = \frac{2(1-2P_1)}{(1-2P_1)(W_1+1) + W_1P_1(1-(2P_1)^{m_1})},$$
 (6)

$$\tau_2 = \frac{2(1-2P_2)}{(1-2P_2)(W_2+1) + W_2P_2(1-(2P_2)^{m_2})},$$
 (7)

$$P_1 = 1 - (1 - \tau_1)^{c-1} (1 - \tau_2)^{N-c},$$
(8)

$$P_2 = 1 - (1 - \tau_1)^c (1 - \tau_2)^{N - c - 1},$$
(9)

where *c* is the number of nodes with real time traffic. Then, τ_1 , τ_2 , P_1 , and P_2 can be solved correspondingly. Once these parameters are known, we can obtain P_{tr} , P_{a1} , P_{a2} , and P_{s} :

$$P_{\rm tr} = 1 - (1 - \tau_1)^c (1 - \tau_2)^{N-c}, \tag{10}$$

$$P_{\rm al} = \frac{\tau_1 (1 - \tau_1)^{r-1} (1 - \tau_2)^{r-1}}{P_{\rm tr}},\tag{11}$$

$$P_{a2} = \frac{\tau_2 (1 - \tau_1)^c (1 - \tau_2)^{N - c - 1}}{P_{\text{tr}}},$$
(12)

$$P_{s} = \frac{1}{P_{tr}} \Big(c\tau_{1} (1 - \tau_{1})^{c-1} (1 - \tau_{2})^{N-c} + (N - c)\tau_{2} (1 - \tau_{1})^{c} (1 - \tau_{2})^{N-c-1} \Big).$$
(13)

4.2. Delay analysis

The average delay time for one hop equals to the average number of retransmission multiplied by the average renewal cycle time plus the successful transmission time [16] and [17]. The detailed parameters are listed below:

- *E*[*P*₁], the average length of the longest real-time packet payload;
- *E*[*P*₂], the average length of the longest non-real-time packet payload;
- *E*[Idle], the number of consecutive idle slots between two consecutive transmissions on the channel

$$E[\text{Idle}] = \frac{1}{P_{\text{tr}}} - 1; \qquad (14)$$

- *H*, the packet header size in IEEE 802.11;
- δ , the propagation delay;
- *T*_c, the average time the channel is sensed busy by the stations during a collision;

•
$$T_{s1} = RTS + SIFS + \delta + CTS + SIFS + \delta + H + E[P_1] + SIFS + \delta + ACK + DIFS + \delta;$$

• $T_{s2} = RTS + SIFS + \delta + CTS + SIFS + \delta + H + E[P_2] + SIFS + \delta + ACK + DIFS + \delta;$

•
$$T_{\rm c} = RTS + DIFS + \delta$$

We define the average delay as follows:

$$D_1 = R_1 (T_{\rm f1} + E[\rm idle]) + T_{\rm s1},$$
(15)

$$R_1 = \frac{1}{P_{a1}} - 1,\tag{16}$$

$$T_{\rm f1} = \frac{P_{\rm s} - P_{\rm a1}}{1 - P_{\rm a1}} \left(\frac{c P_{\rm a1}}{c P_{\rm a1} + (N - c) P_{\rm a2}} T_{\rm s1} + \frac{(N - c) P_{\rm a2}}{c P_{\rm a1} + (N - c) P_{\rm a2}} T_{\rm s2} \right) + \frac{1 - P_{\rm s}}{1 - P_{\rm a1}} T_{\rm c},$$
(17)

where D_1 is the average real-time packet delay, R_1 is the average number of required re-sensing channel and T_{f1} is the expected time between two consecutive channel sensing.

For non-real-time packets, the corresponding equations can be described as follows:

$$D_2 = R_2 (T_{\rm f2} + E[\rm{idle}]) + T_{\rm s2},$$
(18)

$$R_1 = \frac{1}{P_{a2}} - 1,\tag{19}$$

$$T_{f2} = \frac{P_{s} - P_{a2}}{1 - P_{a2}} \left(\frac{cP_{a1}}{cP_{a1} + (N - c)P_{a2}} T_{s1} + \frac{(N - c)P_{a2}}{cP_{a1} + (N - c)P_{a2}} T_{s2} \right) + \frac{1 - P_{s}}{1 - P_{a2}} T_{c},$$
(20)

where D_2 is the average real-time packet delay.

Based on the equations above, we can further derive the end-to-end delay. We analyze the system by a *k*-stage Erlangian server, in which we take *k* as the number of hops required for a packet to reach the destination, and μ as one hop delay. Then, we can define the end-to-end delay distribution in this system as follows:

$$b(x) = \frac{\mu(\mu x)^{k-1} e^{-\mu x}}{(k-1)!}$$
(21)

Table 2 The parameters of MAC for simulation.

1	
Attribute	Value
Channel bandwidth	2 MB
DSSS preamble	144 bits
DSS header	48 bits
RTS size	176 bits
CTS size	128 bits
ACK size	112 bits
SIFS time	$10 \ \mu s$
DIFS time	$20 \ \mu s$
Slot time	$20 \ \mu s$
Retry limit	6 time

and we calculate the value of the peak:

$$x_{\text{peak}} = \frac{k-1}{\mu},\tag{22}$$

$$k = hc + 1, \tag{23}$$

where k is the number of nodes and hop count (hc) will be k - 1. As $\mu = 1/D_1$ is adopted, the result determines the delay of a real-time packet going through WLAN. If $\mu = 1/D_2$ is adopted, the result determines the delay of a non-real-time packet.

Furthermore, we extend the analysis of [18] to estimate the delay of access delay of cellular system. The probability to have u access request messages during a period of T slots, reserved as PRACH slots, is given by the following equation:

$$P(T, u) = \frac{(\lambda T \Delta s)}{u!} e^{-\lambda T \Delta s}, \qquad (24)$$

where Δs is the period of one PRACH slot, and λ is defined as

$$\lambda = \frac{1}{\frac{P_{a1}^2}{P_{a1} + P_{a2}} + \frac{P_{a2}^2}{P_{a1} + P_{a2}}}.$$
 (25)

Then we can define the probability of a successful access request, P_{suc} , and the average access delay, D_{RA} , as follows:

$$P_{\rm suc} = \sum_{u=1}^{\infty} P(T, u) \cdot (1/T) \cdot \left((T-1)/T \right)^{u-1}, \quad (26)$$

$$D_{\rm RA} = \frac{1}{P_{\rm suc}} \left(\frac{T}{2} \cdot \Delta s \right). \tag{27}$$

According to these equations, the end-to-end delay through cellular system can be defined as:

$$2WD + D_{RA} + 2 \cdot \frac{\text{average length of payload}}{\text{cellular transmission rate}}, \qquad (28)$$

where WD is the delay from the source node to a node with cellular device. As $\mu = 1/D_1$ is adopted, the delay of a realtime packet through cellular system will be determined. As $\mu = 1/D_2$ is adopted, the delay of a non-real-time packet will be determined.



Figure 13. End-to-end delay distribution of real-time packet.



Figure 14. End-to-end delay distribution of non-real-time packet.



Figure 15. P_{a1} and P_{a2} versus various traffic combinations with distinct initial backoff window.

4.3. Numerical results

The numerical results have been obtained using the system parameters listed in table 2. Figures 13 and 14 show the endto-end delay distributions of a real time packet and non-realtime packet versus various hop count. Figure 15 shows P_{a1} and P_{a2} versus various traffic combinations with distinct initial backoff window. We can see that in figure 16 demonstrates when w1 = 8 and w2 = 128, the real time pack-



Figure 17. Connectivity versus node number.

ets have (20 times) the much higher probability of successful transmissions.

The theoretical analysis of the probability of successful transmission demonstrates that the voice traffic always has higher priority to access medium. It proves that the scheme we propose can guarantee the QoS for real time packets.

5. Simulation and performance analysis results

Our first objective is to show the connectivity between nodes for various network conditions. The default values used in this simulation are shown in table 2. Individual data traffic follows Poisson distribution model and each real-time connection transmits 8 kbps voice data per second. The network consists of a predefined number of mobile nodes in a 12000×12000 square meter area. Radio transmission range is 300 meters for WLAN devices and 2300 meters for cellular devices; radio condition is assumed to be ideal that there is no variable SNR (signal to noise ratio). The parameter of "cellular active percentage" denotes the percentage of cellular active subscribers to the whole subscribers. For example, 20% cellular active percentage of 100 nodes denotes that 20 nodes are serviced and active, and the others are inactive. Connectivity might not sustain, either because there are considerable subscribers in the same cell for the base station to handle or because particular subscribers are out of the cell coverage for the base stations to reach and service. The parameter of "con-



Figure 18. Number of received packets versus node number (data).



Figure 19. Number of received packets versus node number (voice).

nectivity" is defined as that the number of connections between any two nodes. For example, the complete connectivity is N(N-1)/2 for N nodes in the system.

For the cellular network, the connectivity of higher cellular active percentage is undoubtedly is more than that of lower cellular active percentage (figure 17). HWN has the similar trend. However, the connectivity in HWN is greater than that in cellular network, especially when the number of node is larger and the cellular service percentage is lower. This is because nodes not serviced in a cell or outside of any cell coverage can use HWN routing path to access nodes with cellular device and communicate with another nodes far away. Therefore, when the number of node is large enough, HWN still can achieve the whole connectivity regardless of cellular service percentage. But connectivity is completely based on cellular service percentage. On the other side, connectivity in HWN and cellular system is much larger than that in an ad hoc network, so we do not show the connectivity in ad hoc network.

Operations of the proposed routing protocol are simulated, including the process of register, routing path discovery, and data/voice transmission. 100, 200, 300, 400 and 500 nodes are placed randomly in the environment to study the impact of QoS support. The scheme of QoS support will guarantee the delay of real-time packets.

In figures 18 and 19, the results not only show the total number of received data and voice packets in HWN is larger



Figure 20. Delay of packets versus hop number (data).



Figure 21. Delay of packets versus hop number (voice).

than that in cellular network, but also show the number of received voice packets with QoS support that is larger than without QoS support in HWN. However, QoS support scheme does not influence data transmission significantly.

In figures 20 and 21, we show the delay of packets on different hop number, different number of nodes with/without QoS support. Without QoS support for voice transmission, delay time is usually various, so the curves are unsettled. But with QoS support, the delay time is reduced and almost no more than 10 ms on 6 hops. Similarly, the delay of data packets with or without QoS support is analogous and is not influenced. The results in figure 21 indicate that delay on fewer nodes is smaller in HWN. Through these simulations, we can confirm that the proposed QoS support can guarantee good quality real-time traffic through bandwidth reservation of connection, higher priority and smaller backoff tiime, and queue reservation.

For detailed study, in figures 22 and 23, the results exhibit the delay time related to different max hop limitation. Higher max hop limitation usually causes more packets that are broadcasted. And the higher max hot limitation the more probability of collision. Without QoS support, the delay time increases rapidly and unstably as the max hop limitation increases. Nevertheless, with QoS support, the delay time rises slowly and stably. It indicates that voice packets are always sent on time on one or more hops.



Figure 22. Delay of packets versus max hop limitation (data).



Figure 23. Delay of packets versus max hop limitation (voice).



Figure 24. Number of received packets versus cellular active percentage.

The parameter of cellular active percentage influences the number of packets received (or sent) (figure 24). As the cellular active percentage gets high, there are extra luxuries to access cellular network due to higher connectivity and then there are more path options for the packets to cross a cellular network and an ad hoc network. Therefore the number is increasing with the increase of cellular active percentage.

The affordable bandwidth will be improved to higher range in the future cellular networks, such that GPRS offers 100 kb/s and 3G [18] offers 2Mb/s. Furthermore, received packets are evaluated depending on different provided band-





width of cellular devices in HWN and the results are in figure 25. In this figure, the number of packets received is increasing quickly with the increasing of provided bandwidth unquestionably, although the increasing rate of the number is not the direct ratio to the increasing rate of provided bandwidth.

6. Conclusion and future work

This paper proposes a new architecture: Heterogeneous Wireless Network (HWN), networking ad hoc network and cellular network, and presents a dynamic adaptive routing protocol (DARP), a novel routing protocol for the integrated system. HWN provides mobile users versatile communication with anyone or any device at anyplace and anytime, and DARP aims to provide a routing path with appropriate QoS guaranteed for this hybrid network. Four essential factors are introduced in the protocol. First, the routing process facilitates nodes to construct an optimal way toward the destination via a cellular network or an ad hoc network or both. Second, the register process assists nodes to attach cellular network for further contact by others. Third, the routing table maintenance coordinates to avoid the waste of bandwidth and reconstructs from broken paths. Fourth, QoS routing guarantees the delay and bandwidth for real-time traffic connections.

The simulation results demonstrate the connections accommodated by HWN system are more than either those in a cellular network or in an ad hoc network. Besides, the ratio of connections in HWN system to those in cellular network increases as radio range of cellular devices decreases. What's more, the number of the routing paths in HWN also increases as the number of nodes in the system increases. As a result of avoiding sending same packets twice, routing table in HWN maintains the performance, and routing packets will not increase as quickly as the radio range of cellular devices decreases.

The mobility of nodes usually causes connection failure. Although we propose a process of routing path maintenance structure, the detail influence of low mobility and high mobility on routing traffic load and delay may be further considered. Future research efforts could extend DARP protocol to suit different speeds of mobile nodes. Another great challenge is to guarantee quality of service to a mobile node. HWN system would provide a new path and reserve required bandwidth on the new path, and also release the bandwidth used on the old path.

Besides, with the technology advances of wireless communications and increasing needs of multimedia connections, a great number of wireless devices will need to be upgraded to fit higher bandwidth and lower cost. The typical new devices include improved high-speed Wireless LAN, GPRS (General Packet Radio Service), Bluetooth, and 3G. The improved high-speed Wireless LAN can provide up to 11 Mb/s or 20 Mb/s, and the GPRS can offer up to hundreds kb per second. Moreover, 3G can support up to 2 Mb/s. Hence, those devices will offer various bandwidth options in a cellular network or an ad hoc network, and then a optimal routing protocol will be constructed for these changes to achieve better performance.

References

- J.E. Padgett, C.G. Gunther and T. Hattori, Overview of wireless personal communications, IEEE Communications Magazine 33(1) (1995) 28–41.
- [2] IEEE Standards Board, Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications, IEEE Std 802.11-1997, The Institute of Electrical and Electronics Engineers, Inc. (1997).
- [3] C. Hamling, An overview of GSM data services, in: Proc. of IEE Colloquium on Roaming with Data (1995) pp. 2/1–2/8.
- [4] D.J. Goodman, Trends in cellular and cordless communication, IEEE Communications Magazine 29 (June 1991) 31–40.
- [5] R. Iyer, J. Parker and P. Sood, Intelligent networking for digital cellular systems and the wireless world, in: *Proc. of IEEE GLOBECOM*, Vol. 1 (1990) pp. 475–479.
- [6] S.G. Niri and R. Tafazolli, Cordless cellular network integration for the 3rd generation personal communication systems, in: *Proc. of 48th IEEE VTC 98*, Vol. 1 (1998) pp. 402–408.
- [7] M. Albrecht, M. Frank, P. Martini, M. Schetelig, A. Vilavaara and A. Wenzel, IP services over Bluetooth: leading the way to a new mobility, in: *Proc. of Local Computer Networks*, LCN (1999) pp. 2–11.
- [8] J.C. Haartsen, The Bluetooth radio system, IEEE Personal Communications 71 (February 2000) 28–36.
- [9] A.T. Campbell, J. Gomez and A.G. Valko, An overview of cellular IP, in: *Proc. of IEEE WCNC*, Vol. 2 (1999) pp. 606–610.
- [10] Y.-D. Lin and Y.-C. Hsu, Multihop cellular: a new architecture for wireless communications, in: *Proc. of IEEE INFOCOM 2000*, Vol. 3 (2000) pp. 1273–1282.
- [11] J. Leskela, Deterministic multihop radio on the top of 802.11 MAC, in: Proc. of 10th Euromicro Workshop on Real-Time Systems (1998) pp. 71–78.
- [12] Y.-D. Lin, Y.-C. Hsu, K.-W. Oyang, T.-C. Tsai and D.-S. Yang, Multihop wireless IEEE 802.11 LANs: a prototype implementation, in: *Proc. of IEEE ICC*, Vol. 3 (1999) pp. 1568–1572.
- [13] S.R. Das, R. Castaneda, Y. Jiangtao and R. Sengupta, Comparative performance evaluation of routing protocols for mobile, ad hoc networks, in: *Proc. of 7th International Conference on Computer Communications and Networks* (1998) pp. 153–161.
- [14] D. Johnson and D. Maltz, Dynamic source routing in ad hoc wireless networks, in: *Mobile Computing*, eds. T. Imiclinski and H. Korth (Kluwer Academic, 1996).
- [15] C.E. Perkins and E.M. Royer, Ad-hoc on-demand distance vector routing, in: Proc. of WMCSA 1999, 2nd IEEE Workshop, pp. 90–100.
- [16] G. Bianchi, IEEE 802.11-saturation throughput analysis, IEEE Communications Letters 2(12) (1998) 318–320.

- [17] P.-H. Chuang, H.-K. Wu and M.-K. Liao, Dynamic QoS allocation for multimedia ad hoc wireless networks, in: *Proc. of 8th International Conference on Computer Communications and Networks* (1999) pp. 480–485.
- [18] V.A. Chitre and J.N. Daigle, IP traffic over GPRS-an Internet service oriented analysis, in: *Proc. of IEEE WCNC*, Vol. 3 (1999) pp. 1263– 1267.
- [19] N. Padovan, M. Ryan and L. Godara, An overview of third generation mobile communications systems, in: *IMT-2000, Proc. of IEEE TEN-CON 1998*, Vol. 2 (1998) pp. 360–364.



Eric Hsiao-Kuang Wu received his B.S. degree in computer science and information engineering from National Taiwan University in 1989. He received his Master and Ph.D. in computer science from University of California, Los Angeles (UCLA) in 1993 and 1997. He is an Associate Professor of Computer Science and Information Engineering at National Central University, Taiwan. His primary research interests include wireless networks, mobile computing and broadband networks. He is a member of IICM

(Institute of Information and Computing Machinery) and IEEE. E-mail: hsiao@csie.ncu.edu.tw



Yi-Zhan Huang received his B.S. degree in computer science and information engineering from National Central University in 1999. He is a graduate student of Computer Science and Information Engineering at National Central University, Taiwan. His primary research interests include wireless networks and mobile ad hoc networks.

E-mail: cfox@nw58.csie.ncu.edu.tw