Available online at www.sciencedirect.com
ScienceDirect
The Journal of Systems and Software 81 (2008) 1625–1639
The Journal of Systems and Software
www.elsevier.com/locate/jss

An agent-based metric for quality of services over wireless networks

Yaw-Chung Chen *, Wen-Yen Chen

Department of Computer Science, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 300, Taiwan

Available online 26 February 2008

Abstract

In a wireless LAN environment, clients tend to associate with the nearest access point (AP) which usually provides the strongest signal. However, this does not guarantee that users will receive the best quality of service (QoS) if the population sharing the network capacity were not considered. In other words, within the same access point, the more the population is, the less bandwidth each user will share, and the worse the quality of service will be. In this paper, we proposed an anticipative agent assistance (AAA) which is an agent-based metric for evaluating and managing the resource information of the wireless access points, computing the potential AP list, and providing clients with resource information of APs. We also propose a novel QoS feedback mechanism which allows users to promptly adjust the service quality with AAA according to the throughput and delay requirements. We evaluate the performance of our proposed method using the ns-2 simulator. Numerical results show that AAA achieves: (1) reduce the transmission delay, (2) increase the throughput, (3) improve the network utilization, (4) accommodate more users to access the networks, and (5) achieve load-balancing. Our metric is implementation feasible in various IEEE WLAN environments.

Keywords: Agent-based metric; Wireless LAN; Quality of services; Anticipative agent assistance; QoS feedback

1. Introduction

With an unprecedented growth in the number of wireless users, applications, and communication technologies, wireless local area networks (WLANs) based on IEEE 802.11 (IEEE 802.11b/d3.0) have become a popular solution for providing network services in campuses, corporate complex, and public places. Some key challenges to deploy cost effective wireless LANs are regarding capacity planning, load-balancing, and resource utilization to provide satisfactory quality of services (QoS) to users.

According to the recent studies on the analysis of wireless networks (Tang et al., 1999, 2000, Kotz et al., 2002), the user population is often distributed highly uneven among wireless access points (APs). Users tend to be situated in particular areas for several reasons, such as the network accessibility, the location of power outlets, and personal fondness. This generally results in heavy traffic in certain areas, and deteriorating the quality of service. To address this problem, we proposed so-called anticipative agent assistance (AAA) to improve quality of service (QoS) over wireless networks. The goals of our approach are (1) to reduce the transmission delay, (2) to increase the throughput, (3) to improve overall network utilization, (4) to accommodate more users to access the network, and (5) to provide load-balancing mechanism. In addition to AAA, we also proposed a novel QoS feedback mechanism, which allows users to dynamically adjust the service quality with AAA according to the throughput and delay requirements. We evaluate the benefit of the AAA system using the ns-2 simulator (The Network Simulator), and the numerical results show that our approach performs well under various user configurations, and we conclude that quality of service over wireless networks could be significantly improved through the implementation of our proposed system.

The rest of this paper is organized as follows. Section 2 introduces the related works. The design of AAA system is discussed in Section 3. In Section 4 we evaluate the performance benefits of our approach via simulation. Finally,

* Corresponding author.
E-mail addresses: ycchen@csie.nctu.edu.tw (Y.-C. Chen), wychen@csie.nctu.edu.tw (W.-Y. Chen).

0164-1212/$ - see front matter © 2008 Published by Elsevier Inc.
doi:10.1016/j.jss.2008.02.025

we conclude the paper and address the future work in Section 5.

2. Related works

There have been a number of proposals to improve the quality of service (QoS) in WLANs (IEEE 802.11a/b/g), including the media access control (MAC) layer enhancement, network-layer support, system approaches, and hardware solutions. We will further discuss these topics in the following sections.

2.1. MAC-layer QoS support

The original 802.11 medium access control (MAC) protocol (IEEE 802.11 Wireless LAN) has defined two major access methods for wireless stations. The distributed coordination function (DCF) which is based on carrier sense multiple access with collision avoidance (CSMA/CA), and the point coordination function (PCF), which is contention-free and aims to support real-time traffic flows (Lindgren et al., 2003).

The DCF is the fundamental access mechanism used to support asynchronous data transfer on a best-effort basis. As identified in the specification, all stations must support the DCF, for use within both ad hoc and infrastructure modes. However, it has the following limitations: (1) only support best-effort service, (2) no guarantee in bandwidth, packet delay and jitter, and (3) throughput degradation under the heavy load. The PCF is an optional access method that intends to support time-bounded services, it provides contention-free frame (CF) transfer and is only usable on infrastructure network configurations. The operation is essentially polling, with the PC performing the role of the polling master. PCF also has some constraints: (1) inefficient and complex central polling scheme, (2) unpredictable beacon delay due to incompatible cooperation between contention period (CP) and contention-free period (CFP) modes, (3) transmission time of the polled stations is undeterministic, and (4) PCF implementation is hard to find. The DCF and the PCF may coexist in a manner that allows both to operate concurrently within the same BSS. When a PC is operating in a BSS, the two access methods alternate, with a contention-free period followed by a contention period. Nevertheless, problems associated with the coexistence of DCF and PCF are addressed in (Grow et al., 1997, Veeraraghavan and Cocker, 2001, Visser and ElZarki, 1995), and especially PCF performs poorly in the existence of DCF.

To support QoS better, IEEE 802.11e (Mangold et al., 2002) defines enhancements, enhanced distributed coordination function (EDCF) and hybrid coordination function (HCF), to 802.11 MAC. The EDCF in 802.11e is the basis for the HCF, like the DCF in 802.11 provides services for PCF. EDCF defines the traffic categories (TCs) for supporting traffic differentiation, and each station has eight TCs, or priority levels. The EDCF provides differentiated services based on traffic categories, but the relative performance is not easy to control because the performance is not proportional to the backoff factor ratios, instead it depends on the number of contending stations. While the HCF is designed to improve polling mechanism of PCF, and provides policing and deterministic channel access by controlling the channel through the hybrid coordinator (HC). The HCF can operate in CFP and CP, and support both InterServ and DiffServ.

Within the HCF, there are two access mechanisms called the enhanced distributed channel access (EDCA) and HCF controlled channel access (HCCA). The EDCA is an extension for 802.11 DCF, and supports 8 priorities (traffic classes). With the admission control in EDCA, the contention window and backoff times are adjusted to change the probability of gaining medium access to favor higher priority classes. Fig. 1 shows the priority values (0–7) which is identical to the IEEE 802.1D priorities. The HCCA uses a HC to centrally manage medium access. Under control of the HC, a nearly continuous sequence of frame exchanges can be maintained, with short, fixed delays between frames. Moreover, HCCA guarantees strict QoS for specific flows from applications with individual QoS parameter.

In addition to these mechanisms proposed in IEEE 802.11e, there has been a number of proposals for MAC protocol in WLANs to support differentiated services (Ada and Castellucia, 2001, Bianchi et al., 1996, Barry et al., 2001, Kim and Jennifer, 2003, Vaidya et al., 2000). However, all of these methods focus on scheduling and fairness on individual wireless MAC without considering the dynamic nature of various wireless network services.

2.2. Network-layer QoS support

Past research in routing for wireless network, such as ad hoc networking, focus on providing QoS routes and efficient routing. In (Tsai et al., 1997), the authors present a cluster-based approach which can inform the source regarding the bandwidth and QoS available to any destination in the wireless network. This enables the establishment of QoS connections within the wireless network and the efficient support of real time, multimedia traffic. In Chen and Nahrstedt (1999), author proposed a distributed QoS routing scheme which selects a route with sufficient resources to satisfy certain delay (or bandwidth) requirement in a dynamic multi-hop mobile environment. These algorithms also work with imprecise state information to select the best qualified one among multiple paths.

In mobile wireless network, the locations and resources of mobile nodes vary dynamically. The state information used in the traditional ad hoc routing protocol (DSDV, DSR, and AODV) may be obsolete very soon due to node mobility. A predictive location-based QoS routing scheme (Shah and Nahrstedt, 2002) is proposed based on a low-cost location-resource update protocol to assist predicting future routes before existing routes break and avoiding route re-computation delay. In addition to the research on QoS routing, there have been some works done on load-balancing routing for

Fig. 1. EDCA supports 8 priority values (traffic classes).

wireless access networks. A new distributed routing algorithm (Hsiao et al., 2001) performs the dynamic load-balancing for wireless access networks. The algorithm introduced the idea of load-balanced backbone tree, which simplifies routing and avoids per-destination state for routing and per-flow state for QoS reservations.

The aforementioned research mainly focus on multi-hop wireless networks, thus each node has to find a QoS route to egress node via connecting to multiple nodes within the network. However, our proposed method will focus on infrastructure wireless network, which means each node in the wireless network will directly connect to the nodes through an access point (one hop).

2.3. Hardware-based QoS support

Recently, wireless vendors like Cisco (Cisco Aironet 350 Series) and OriNOCO (Agere Systems, 2002, Boingo Software, 2002) have their wireless client adapters equipped with site-survey tools to detect the signals from the neighbor access points. With this tool users can learn statistical information regarding the access points, including nearby access points, signal strength and quality, receive and transmit statistics, and link status. As to the data gathered by this tool, users can decide which access point to associate with for the future communication. For example, Fig. 2 illustrates the available sites detected by OriNOCO wireless PC card. From the figure listed below, we learn that there are three available access points surrounding us: CC201, default, and NCTU. Each of them have the equal signal strength, but we choose to associate with AP CC201 as suggested by Boingo software.

In Figs. 3 and 4, Cisco's Aironet Client Utility (ACU) provides more information on signal strength and link quality. From the figures below we have been informed of the channel, network type, power save mode, AP MAC address, etc. Moreover, we present the delivery and transmission statistics in Fig. 5, and the result of link test in Fig. 6.

To some extent these information may be helpful for users while selecting access points. In most cases, mobile users will eventually choose the AP with stronger signal strength for better link quality. However, this does not guarantee better quality of service as there might be many users access the same AP simultaneously. Moreover, it is a burdensome work for wireless users to keep tracking the status of each AP during the connection. Also, these

Fig. 2. Available signals by Boingo.

Fig. 3. Site survey by Cisco Aironet client utility.

Fig. 4. Aironet client adapter status.

Fig. 5. Receive and transmit statistics.

utilities only inform the users of surrounding APs rather than distant and least-loaded APs. For instance, a traveler in the airport is waiting for his flight to LA and is situated at gate A1. Since LA is a popular scenic spots, you can imagine that there will be much more people located in A1 gate rather than other gates. B1, another airport gate, is about two or three corners next to A1 and with fewer passengers than A1. However, the site-survey tool only notifies the traveler about the APs installed in the ceiling of A1, A2, or nearby gates but cannot detect the APs of B1 due to the longer distance and physical obstructions (people, stores, and walls). In our proposed system we are trying to solve this problem.

2.4. System-based QoS support

Microsoft Corp. has incorporated a new feature called Microsoft Wireless Configuration Manager (MWCM) in

Fig. 6. Link test by ACU.

its operating systems (Windows 2000/XP). MWCM helps users configure the settings of wireless network such as the switching between infrastructure and ad hoc modes, security option (Wired Equivalent Privacy, WEP), and 802.1x authentication support (Extensible Authentication Protocol, EAP; Remote Authentication Dial In User Service, RADIUS).

In Zerfos et al. (2003), the authors introduce DIRAC, a software-based router system that is designed for wireless networks to support three wireless network services: link-layer assisted fast handoff, channel-adaptive scheduling, and link-layer enforced policing. DIRAC is a distributed router architecture that consists of a router core (RC), and several router agents (RAs) at each access point/base station. When a mobile user roams from an old access router to a new one, link-layer handoff will happen first. With the DIRAC system framework support, it uses the *ReAssociation* message of link-layer handoff process that precedes the network-layer, and uses it as the event that triggers the network-layer handoff service. Fig. 7 shows the message flow of the protocol.

Fig. 7. Link-layer informed fast handoff.

Although DIRAC enables wireless protocol and provides router services for wireless networks, it is unable to inform the mobile users ahead of moving to the access point or base station with more available resources. Hence, we will take this feature into account in our system design, and here we define the action of informing the mobile users beforehand as anticipative assistance.

3. Proposed approaches

3.1. Motivation

In recent years, the deployment of wireless services has greatly increased. Through the discussion in Section 2, we have learned that wireless devices and utilities tend to sense the stronger signal and to associate with the nearest access point by default. Besides, we usually consider that the best link quality guarantees the best quality of service. However, a user may find an opposite situation because the service quality may still be poor even you choose the 'qualified' AP. The following two scenarios are demonstrated as examples.

Fig. 8 shows four access points installed in a conference room. Here we assume there might be physical obstructions such as walls or pillars within the wireless coverage of each access point. While you enter into the room, the wireless utilities will soon notify you of the AP-1 as your best choice since AP-1 features strongest signal near the door. Nevertheless, it is very likely that service provided by AP-1 is worse than that provided by AP-4 because the former should serve much more users.

In Fig. 9, you are located around G4 in a busy international airport waiting for the flight to San Francisco (SF). The departure time of the flight is 15 min later and it starts on board already. Since SF is a hot business destination, G4 is crowded and many passengers get on line through the wireless link. At this moment, your boss phones you and asks you to send some important documents to him via email immediately. You promptly attempt to send the email, but you may find the link speed is too slow to make it through by the time of taking off. Meanwhile, G7, another gate three corners away from G4 has much fewer wireless users. Unfortunately, you were not aware of the situation. Hence, the key missing piece is a system framework that is able to inform you the overall status of the network environment, and provides assistance beforehand.

Fig. 8. Scenario I.

Fig. 9. Scenario II.

Based on the above ideas, we propose a novel approach called anticipative agent assistance (AAA), which is a centralized framework providing various forms of interaction between users and agents in order to improve the quality of service over wireless networks.

3.2. Anticipative agent assistance

3.2.1. Design rationale of AAA

AAA is designed for supporting users to acquire better quality of service through interaction with agents. Basically, AAA is a centralized architecture which is composed of a sever agent (SA) and numerous client agents (CAs) installed at each access point/base station. Fig. 10 illustrates the AAA architecture.

Fig. 10. AAA architecture.

Here the connection between a SA and a CA is wired, and a CA at access point/base station provides wireless coverage for the mobile users. However, the connection between SA and CA is not necessary through wired link, it could be through wireless connection. Since the communications protocol between SA and CA is TCP/IP based, so it makes no difference using wired or wireless link in guaranteed delivery except the signal transmission quality and link bandwidth. To prevent SA from failure which may crash the entire AAA system, we need a back-up mechanism for SA. While each CA can arbitrarily join or leave an AAA group. For example, when a mobile user enters a wireless domain under CA-1 and requests AAA support, CA-1 will generate an assistance request message and forward it to SA. SA refers to its database and control algorithm to calculate the qualified CA list and sends this information to the user through CA-1. With AAA enabled, each mobile user will soon be advised whether to move or not before starting to work on line. The number of CAs supported is not restricted, and it can be applied to different subnets to provide extensive services. Moreover, AAA adopts the centralized design due to the following concerns: (1) Make it easy to control and monitor the entire network status, (2) Lighten the load of access points (CAs), and (3) Reduce the overhead caused by intercommunication among CAs. In the following sections, we describe three forms of interaction in AAA system, including SA–CA interaction, CA–MN interaction, and SA–CA–MN interaction.

3.2.2. SA–CA interaction design rationale of AAA

Two procedures, registration and deregistration are defined in the SA–CA interaction (Fig. 11). The former is initialized by a new CA when it wants to join an AAA group. The CA starts the registration process by sending a *CAReg* message to SA. The *CAReg* message contains the essential information such as CA's ID (MAC address of CA), CA's IP address, $X/Y/Z$ coordinates, number of mobile users, and available bandwidth. Once the registration requested by a CA is approved, a *CAAck* message initialized by SA will be sent back to the CA to finish the registration procedure. Otherwise, a *CAFail* message with failure code will be sent to CA. Regarding the latter, a CA can quit the AAA group by sending a *CADReg* message to SA for reminding SA to remove its record from CA-record (Table 1). SA maintains a CA-record in its database as reference and input for future assistance queries from all CAs. All necessary information is stored in SA rather than CAs, and SA also takes charge of the mutual communication among CAs. Thus, CA is exempted from maintaining an individual database and frequent communications with other CAs. A lightweight CA has been fulfilled via the design of centralized architecture.

Table 1
CA-record maintained in SA

ID	IP	Loc X	Loc Y	Loc Z	MN	Bw
CA1	100.0	50.0	50.0	0.0	5	11Mb
CA2	200.0	550.0	50.0	0.0	3	54Mb
CA3	300.0	550.0	550.0	0.0	8	11Mb
CA4	400.0	50.0	550.0	0.0	1	54Mb

3.2.3. CA–MN interaction

In the interaction between CA and MN (mobile node or user), there are three kinds of procedures (Fig. 12). The first one is advertisement procedure. Each CA at an access point/base station periodically broadcasts advertisement messages. For the overall network performance concern the frequency of advertisement broadcasting can be set to once per minute. The *CAAdv* message is composed of CA's ID, CA's IP address, elapsed time, and AAA-enforcement option. Elapsed time field in *CAAdv* message defines the valid lifetime of an advertisement, and AAA-enforcement option (ON/OFF) is to determine whether to ask all mobile users to enable AAA support for the purpose of load-balance. The default value of AAA-enforcement option is OFF, that means we allow mobile users to decide whether to enable AAA support in most cases.

Upon receiving the advertisements from CA, MN will soon check its AAA setting. If the AAA is configured as ON, MN will send a *ProbeReq* message to CA for requesting AAA services (detailed discussion in next section); otherwise MN acts as a normal node in the wireless network. After that, MN may move to another wireless domain of a new CA or stay with the current one. Then MN has to initiate a registration process which sends a *MNReg* message to the corresponding CA, and the content of message includes MN's ID, MN's IP address, AAA support, and QoS option. The value of AAA support is based

Fig. 11. SA–CA interaction.

Fig. 12. CA–MN interaction.

on users' AAA setting, and QoS option will be introduced in the SA–CA–MN interaction. While CA accepts registration request from a MN, it will update its information in SA's CA-record and a registration acknowledgement will be sent back to the MN.

The last procedure in CA–MN interaction is deregistration. If a MN intends to leave for another location, it starts a deregistration process which sends a *MNDReg* message to the associated CA and waits for the reply to terminate the procedure. In the meantime, when CA receives the *MNDReg* request from a MN, it updates the data in SA's CA-record.

3.2.4. SA–CA–MN interaction

Since MN has two settings, AAA support and QoS option, the combination of these two values will result in three different types of interaction, [AAA OFF/QoS OFF], [AAA ON/QoS OFF], and [AAA ON/QoS ON] (QoS can be ON only when AAA is ON).

3.2.4.1. Case 1: [AAA OFF/QoS OFF]. If both AAA and QoS are disabled in MN's configuration, MN will behave as an ordinary mobile node in the wireless network, and the AAA system seems never exists. However, there is still an action needed to be done by the MN, the registration process. When a MN enters the coverage of CA-1, it soon hears the advertisement and then sends a *MNReg* message to CA-1. After checking the value of AAA field, CA-1 generates a $CF + MNReg$ message to SA for updating the field and number of mobile users in CA-record. When the $CF + MNReg$ message arrives at SA, SA processes the request and modifies the corresponding records. Once the update has been complete, a $CF + MNAck$ message is created by SA to inform CA-1 of successful modification. Finally, CA-1 sends a *MNAck* message to terminate the registration procedure. The detailed message flow is elaborated in Fig. 13.

3.2.4.2. Case 2: [AAA ON/QoS OFF]. In case 2, the AAA support is turned on and QoS option is turned off. While a MN gets into the wireless coverage of CA-1, it soon senses the advertisement and checks its AAA setting. Since the AAA is ON, the MN sends a *ProbeReq* message to CA-1 right away. The *ProbeReq* message contains MN's ID, MN's IP address, *X*-coordinate, *Y*-coordinate, *Z*-coordinates, and distance-sensitive option, which is used in the *CalBest* algorithm described below and is configured by a MN. CA-1 checks the type of received message and attaches its IP address to the end of message to form a new message, $CF + PbReq$. The $CF + PbReq$ message is forwarded to SA, which recognizes the $CF + PbReq$ message as an AAA query and uses the *CalBest* algorithm to output the recommended movement in the $CF + BstAck$ message.

The *CalBest* algorithm takes four parameters as its input, including *X*-coordinate, *Y*-coordinate, *Z*-coordinate, and distance-sensitive option in the *ProbeReq* message. If the distance-sensitive option is OFF, which means to select a least-load AP without considering the distance between the mobile user and the new AP, *CalBest* sorts the CA-record based on the number of mobile users and available bandwidth. In Table 1, CA-4 will be recommended to the MN in the end. On the other hand, if the distance-sensitive option is ON, *CalBest* algorithm first repeats the sort procedure described above. After that, the qualified finalists (with slighter load than current AP) are sorted again based on the distance to the mobile user. In Table 1, CA-2 replaces CA-4 as the best choice.

Following the output of *CalBest*, SA generates a $CF + BstAck$ message to CA-1. CA-1 informs MN of recommended movement by sending a *BstAck* message. MN moves to the location as the message suggests and initiates the regular registration procedure presented in Section 3.2.3. Fig. 14 illustrates the message flow among SA, CA, and MN.

3.2.4.3. Case 3: [AAA ON/QoS ON]. With the AAA enabled, it guarantees that mobile users get the quality of service in the beginning. In order to keep monitoring and maintaining the quality of the ongoing connection, we introduce the QoS option to the MNs. The purpose of QoS option is to periodically check the throughput, delay or jitter of a connection. Once the value of throughput (delay or jitter) degrades to below a specified threshold for accumulative 60 s, a *ProbeReq* messages will be automatically generated and sent to the corresponding CA on behalf of the MN.

Fig. 13. Message flow in case 1.

Fig. 14. Message flow in case 2.

In case 2, if a MN accepts the suggestion and moves from a heavy traffic area (CA-1) to a light traffic area (CA-4). After getting into CA-4 area, the MN starts the applications and the quality of service is guaranteed in the beginning. As time passed by, there may be more users with AAA off enter into the wireless coverage of CA-4. Nevertheless, the MN was not aware of other users' coming, and the throughput of the connection drop drastically. With the feature of QoS option, a monitor mechanism in MN will spontaneously be enabled and keep tracking the throughput of the connection. If the throughput is under the threshold for accumulative 60 s (user adjustable), the monitor mechanism will send a *ProbeReq* message to CA-4 for another AAA query (same procedure in case 2). Consequently, the MN will soon be advised to move to another area to gain better services. Therefore, the joint use of AAA support and QoS option will ensure the quality of service for the entire connection.

The threshold setting is based on the class of service, and different services have different requirements for throughput, delay, and jitter. Table 2 shows the thresholds for various service classes. Moreover, Fig. 15 demonstrates the message flow in case 3. In Fig. 15, a *ProbeReq* message has been automatically created and sent to CA-4. Meanwhile, MN waits for the *BstAck* message from SA. Once the *BstAck* message is received by a MN, the MN deregisters with CA-4 before moving to the recommended area, CA-2. While the deregistration procedure is complete (MN receives the *MNDAck* message), MN immigrates to CA-2 and then initiates the regular registration process at the same time.

Table 2
Threshold for service classes

Code	Class of Service	Throughput	Delay	Jitter
0	Default	T_D	D_D	J_D
1	Best effort	T_BE	D_BE	J_BE
2	Voice	T_VO	D_VO	J_VO
3	Video	T_VI	D_VI	J_VI
4	Reserved	–	–	–

Fig. 15. Message flow in case 3.

Fig. 16. Solicitation message flow.

3.2.5. Solicitation

In order to make AAA system more flexible, we introduce another feature to the MNs, solicitation function. In the original design, AAA service is requested through sending out a *ProbeReq* message by MNs. However, a *ProbeReq* message only can be generated when the value of AAA support or QoS option is turned on. Namely, this prevents mobile users from asking service whenever needed. With the support of solicitation function, MN is allowed to manually send a *Solicitation* message to request AAA service anytime. Fig. 16 shows the message flow in detail.

3.2.6. Solicitation

In order to make AAA system more flexible, we introduce another feature to the MNs, solicitation function. In the original design, AAA service is requested through sending out a *ProbeReq* message by MNs. However, a *ProbeReq* message only can be generated when the value of AAA support or QoS option is turned on. Namely, this prevents mobile users from asking service whenever needed. With the support of solicitation function, MN is allowed to manually send a *Solicitation* message to request AAA service anytime. Fig. 16 shows the message flow in detail.

When a MN wants to ensure the transmission quality, it could quickly send a *Solicitation* message to request AAA service rather than waits for later reaction initiated by QoS option. In Fig. 16, a MN sends a *Solicitation* message to the currently associated CA, CA-1, and waits for the advices from SA. Meanwhile, MN deregisters with CA-1 so as to migrate to other location. Afterward MN moves to CA-3 and begins the regular registration procedure. Thus, the solicitation function gives MNs the great flexibility supported by our AAA system either automatically or manually.

3.3. Design goals

The AAA system is mainly responsible for managing the control information of the entire wireless access points, outputting the potential AP list with our *CalBest* algorithm, and then ahead informing the users of AP with better service. Through the interactions with agents, MNs are expected to acquire anticipative assistance by our AAA system. Our design goals are (1) to reduce the transmission delay, (2) to increase the throughput, (3) to improve the overall network utilization, (4) to accommodate more users to access the network, and (5) to provide load-balancing mechanism.

Goals (1) and (2) can be easily achieved by the joint use of AAA support and QoS option. With these two features, MNs should be advised to migrate to the area with lighter load and better quality of service. Furthermore, we can achieve Goal (3), (4), and (5) via the AAA-enforcement option in the *CAAdv* advertisement message broadcasted by CAs. Once the AAA-enforcement option is turned on, the system will require all mobile users to be evenly distributed to different areas for the load-balancing consideration. Plus, the AAA-enforcement option is superior to other options. By way of enabling the AAA-enforcement option, it will prevent users from aggregating in some specific areas so as to improve the overall network utilization. In next section, we present the performance evaluation of our proposed approach.

4. Performance evaluation

4.1. Individual user AAA-support enabled

4.1.1. Simulation scenario I

Fig. 17 illustrates the network topology in our first test set. There are four CAs (attached to access points/base stations) separately connecting to the SA with a 100 Mbps/1 ms wired link (the link delay is composed of layer 2 process time and propagation time), and so is the connection between a CN and a SA. 802.11b is used to facilitate the wireless communication for mobile users and CAs. The mobile users start the FTP application destined to the CN with TCP Reno traffic, and here we simulate that a new MN is going to enter CA-1 area. The simulation time is 300 s. Besides, we use various initial user distributions (from 0 to 5) for CA-1 to observe how a new MN benefits from our AAA support. The user distributions for CA-2, CA-3, and CA-4 are 3, 2, and 0, respectively. In order to establish statistical significance of the findings, we have repeated the experiment 5 runs, and the reported numerical results are the average performance over 5 runs.

Fig. 17. Simulation scenario I.

4.1.2. Result and analysis I

Fig. 18 shows the MN's throughput under various user distributions at CA-1. When a MN is entering the wireless coverage of CA-1 where one user is already there, the MN's throughput is about 1079 Kbps. However, the throughput is decreased to 458 Kbps while the number of users reaches 6. We can be sure that the throughput is definitely going down if more users join CA-1 and the MN is not aware of the user population. Namely, the more users aggregate at CA-1, the lower throughput a MN will get. On the other hand, with AAA support, the throughput of MN is improved a lot. MN maintains the throughput at 2158 Kbps no matter how many users are associated with CA-1. This is because MN enables the AAA support so that it is directed by AAA system to move from CA-1 to CA-4, where no other user competes for the resources, before starting the FTP session, thus the MN will surely be better served than at CA-1.

Moreover, the sending rate and RTT are illustrated in Figs. 19 and 20, respectively. With the AAA-support enabled, MN's average RTT is 7.36 ms. From Fig. 19, we can easily recognize that the sending rate with six users is not as stable as that with AAA support. The same situation happens to RTT as well. The RTT with six users ranges

Fig. 18. MN's throughput.

Fig. 19. MN's sending rate.

Fig. 20. Round trip time for MN's FTP application.

Fig. 21. User distribution at each CA.

between 25 ms and 80 ms, but only between 5 ms and 9 ms with AAA support. In addition, we also attempt to add more users to CA-1 (more than 7 users in total). However, the RTT dramatically increases to hundreds milliseconds and the sending rate becomes very unstable. The reason is that the queue size of AP overflows thus seriously deteriorates the quality of service for mobile users. Hence, the probability of aggravating QoS will be substantially decreased by using the AAA system. The user distribution at each CA is summarized in Fig. 21. With the suggestion provided by AAA support, the number of users at CA-1 decreases from 6 to 5, and increases from 0 to 1 at CA-4.

4.1.3. Simulation scenario II

Fig. 22 shows the simulation scenario II. It differs from scenario I in the direction of data transfer, link capacity, and application. In scenario II, CN will be the source that sends the audio flows to the mobile users. The audio flow is 64 Kbps UDP traffic with packet size 210 bytes. While the MNs are locating within the coverage of AP, the CN sends the audio flow to MNs. The link capacities for CN–SA and SA–CA have been changed to 10 Mbps/1 ms, 1 Mbps/10 ms, respectively. The simulation time is still 300 s. We have repeated the experiment 5 runs, and the reported numerical results are the average performance over 5 runs.

4.1.4. Result and analysis II

Fig. 23 shows the delay for audio flow received at a MN. From the figure we learn that the delay is proportional to the number of mobile users. The delay varies from 0.014 s to 0.023 s. If the AAA support is enabled, the delay can be reduced to as low as 0.014 s. We have done simulation 5 times for single user experiment, however, each time the result is very close, so the average is about the same, this is because the simulation scenario is pretty pure (single user join, one variable).

Furthermore, we also attempt to increase the number of user to more than 6. Not surprisingly, the end-to-end delay greatly increases to hundreds of milliseconds (647 ms or higher) and the sending rate becomes very unsteady. The reason is that the queue size of AP overflows thus seriously deteriorates the quality of service.

4.1.5. Consideration for mixed traffic

It is very likely that both FTP and VoIP traffic appear simultaneously, these two traffic classes may compete for the air channel and the delay of VoIP as well as the throughput of FTP will be affected by other traffic type in addition to the similar traffic. Since our metric measures the individual performance parameter such as throughput, delay or jitter, no matter how severe the interference will be, our metric can always measure these performance parameters and advise the new comer to make a better choice. Further, under the IEEE 802.11e environment, FTP and VoIP traffic will be handled based on EDCA (enhanced distributed channel access) and HCCA (HCF controlled channel access) mechanism, respectively. Therefore transmission of these two traffic categories would be

Fig. 22. Simulation scenario II.

Fig. 23. End-to-end delay for audio traffic.

accommodated in well partitioned time periods as long as the total bandwidth utilization is under a certain threshold.

4.2. Individual user AAA support and QoS-option enabled

4.2.1. Simulation scenario III

We use the same network topology (Fig. 17) in scenario III. The initial user distributions are 0, 3, 2, and 0 for CA-1–CA-4, respectively. Since the MN enables both AAA support and QoS option, it is directed to CA-1 as recommended by AAA system (with lease load and shortest distance) in the beginning. In order to verify the benefit introduced by QoS option, we add one mobile user every 30 s at CA-1 until the number of users raised to 6. The QoS threshold is set as 1000 Kbps. The other settings such as link capacity, application, and transfer direction, are identical to those of scenario I. Furthermore, for service continuity, every mobile user has Mobile IP (Perkins, 1996) enabled while migrating to other CAs. The reported numerical results are the average performance over 5 runs.

4.2.2. Result and analysis III

Figs 24 and 25 illustrate the sending rate and RTT for MN's FTP application with or without the support of QoS option. In Fig. 24, the MN's throughput decreases from 2064 Kbps to 1008 Kbps when adding a user to CA-1 at 30th second. At 60th second, another user gets into the CA-1 area (3 mobile users in total), the throughput quickly drops to 712 Kbps. At the same time, the monitor mechanism launched by QoS option keeps tracing the throughput based on the specified threshold for the following 60 s.

At 120th second, the total number of mobile users has reached 5 and MN's throughput keeps decreasing. Thus, the monitor mechanism automatically generates a *Probe-Req* message on behalf of MN to request AAA's support. MN takes the suggestion so as to move from CA-1 to CA-4 at 122nd second. During the handoff (122nd–131st second), MN also starts the Mobile IP to ensure the service continuity. At 131st second, MN arrives at CA-4 and the throughput goes up to 2168 Kbps again. Meanwhile, RTT rapidly decreases from 32 ms to 6 ms. On the contrary, without the QoS option, MN's throughput drops drastically to as low as 336 Kbps and RTT raises to as high as 54 ms.

Fig. 26 summarizes the user distribution at CA-1 for the entire trace. MN's migration (at 122nd second) and new user's participation (at 150th second) results no change in the number of users.

Fig. 24. MN's sending rate.

Fig. 25. Round trip time for MN's FTP application.

Fig. 26. User distribution at CA-1.

4.3. AAA-enforcement option enabled

4.3.1. Simulation scenario IV

Fig. 27 shows the scenario of the load-balancing mechanism in which AAA-enforcement option is enabled through the advertisement messages. The CN connects to SA through a 10 Mbps/1 ms link, and the connection

Fig. 27. Simulation scenario IV.

between SA and each CA is a 2 Mbps/2 ms link. MNs are the senders which initiate the FTP application with TCP Reno traffic, which is destined to CN. In general, users tend to gather in particular area for several reasons, such as the network accessibility, the location of power outlets, and personal fondness. Namely, the mobile users are unevenly distributed at each CA. Therefore, we assume that the user distribution will be 5, 3, 0, and 0 for CA-1–CA-4, respectively, without AAA enforcement. In the following section, we will examine the queue size and queue drop of the links between SA and CAs to see what we benefit from this feature. The default queue size is 50, and the simulation time is 300 s. The reported numerical results are the average performance over 5 runs.

4.3.2. Result and analysis IV

Since the AAA enforcement is enabled in the advertisement messages, every mobile user will be asked to move to the assigned CA while entering the wireless area. In Fig. 28 we can see that all the users are equally dispatched to four CAs. From the network administrator's perspective, this function indeed achieves the load-balancing and reduces the probability of having mobile users congested in certain areas. Furthermore, we in here define the link between SA and each CA (CA-1–CA-4) as Link1 to Link4. Figs. 29–32 show the queue size of each link.

The default queue size of each link is 50 packets. Without the AAA-enforcement enabled, the queue size of Link1 can be very fluctuant and ranges between 0 and 49 (Fig. 29). Once the queue space runs out, it results in packet drop. From Fig. 33 we observe that the packet drop for Link1 can be as high as 960 packets. Link2 experiences the same problem at the same time. The range for queue size of Link2 is between 13 and 49, and also leads to 261 packet drops. Since all users congregate in CA-1 and CA-2, the queue sizes for Link3 and Link4 are always zero, and there is no packet drop. On the other hand, if the AAA enforcement is enabled in the beginning, the packet drop for each link eventually reduces to zero and the queue size of each link keeps steady without overflowing during the entire trace. Figs. 29–32 illustrate that the queue size for each link ranges between 35 and 37. This feature not only increases the possibility of providing quality of service, but also improves the overall network utilization by

Fig. 28. User distribution at each CA.

Fig. 29. Queue size for SA–CA1 link.

Fig. 30. Queue size for SA–CA2 link.

Fig. 31. Queue size for SA–CA3 link.

Fig. 32. Queue size for SA–CA4 link.

Fig. 33. Queue drop for each link.

load-balancing mechanism. In addition to the advantages presented above, this feature may accommodate more users to access the wireless network. Generally speaking, a link with minimum packet drop is critical to real-time services. Time-sensitive services, such as VoIP, video, streaming media, and interactive gaming, require low latency and high network availability. With AAA enforcement, chances are much higher to keep the users staying in the network and to allow more users joining to the same group simultaneously.

We have tried randomly assigning 50/100 users to AAA to see how frequent the movement happened. If the AAA is ON, the average movement per user is 0.86/1.002 for 50/100 users, respectively, during the simulation. Therefore the frequency of the join/leave may not be a significant issue. In case frequent join/leave happened, a high performance AAA server may be required to accommodate the requirement of extensive computation and frequent queries.

4.3.3. Simulation scenario V

The scenario V is shown in Fig. 34. The link between CN and SA remains the same, but the connection between SA and each CA is changed to 1 Mbps/10 ms. CN is the sender which acts as a streaming media server and transmits video stream to mobile users. The video stream is VBR traffic with 384 Kbps sending rate and 800 Byte packet size in UDP. In this scenario, MN will be the first one who enters the CA-1 area. After that, CA-1 takes in one user every 30 s until the total number of users increases up to 8. Moreover, following the assumption in scenario IV, we assume that the user distribution will eventually be 8, 0, 0, and 0 for CA-1–CA-4 without AAA enforcement. In the following section, we can see how the delay and loss rate for MN's application will be affected with the AAA enforcement. The reported numerical results are the average performance over 5 runs.

Fig. 34. Simulation scenario V.

Fig. 35. Delay for MN's video flow.

4.3.4. Result and analysis V

We first investigate the delay and loss rate of MN without AAA enforcement. In Fig. 35, the delay for MN's video flow is kept as low as 19 ms before 60th second. At 61st second, the delay begins to increase and the total number of users at CA-1 has reached 3. Between 61st second and 160th second, three additional users joined CA-1 (6

Fig. 36. Loss rate for MN's video streaming flow.

Fig. 37. User distribution at each CA.

users in total), the delay increases steadily but slowly and the loss rate maintains at zero as shown in Fig. 36. After 160th second, however, both the delay and loss rate suddenly raise and keep going up for the rest of the trace. The 8th user gets into CA-1 area at 210th second, and the delay is as high as 120 ms and loss rate reaches to about 8%. Even there is no other users joining to CA-1 area after 210th second, the delay and loss rate still keeps aggravating in the later half of the trace (210th–300th second). At the end of trace, the delay is raised up to 156 ms and the loss rate climbs up to 22%. On the contrary, if AAA enforcement is enabled, MN and other users will be evenly directed to the four CAs as shown in Fig. 37. Moreover, the MN maintains the delay at 19 ms for the entire trace, and the loss rate is also reduced to 0% till the end. From the aforementioned discussion, we conclude that with the introduction of AAA-enforcement messages, both MN and other mobile users are guaranteed to receive a service at an acceptable level, and meanwhile the goal of load-balancing can be achieved.

5. Conclusions

In this paper, we proposed an agent-based system for improving quality of service over wireless networks. The proposed method, AAA system, is basically a metric mainly responsible for managing the control information of the entire wireless access points, generating the potential AP list with our *CalBest* algorithm, and informing the users of AP (or CA) with least load. Through the interaction with agents, mobile users are expected to acquire anticipative assistance from our AAA system. Furthermore, we introduce a novel idea, the QoS option, which allows users to promptly adjust the service quality with AAA system according to specified threshold of throughput, delay or jitter. For the load-balancing purpose, we incorporate the AAA enforcement option in the advertisement messages. With the AAA-enforcement enabled, all mobile users will be directed by AAA system based on the load-balancing mechanism. Also, we extend the AAA system with the solicitation function. The solicitation feature allows users to request AAA service whenever needed.

We use ns-2 to evaluate the performance of AAA system with various scenarios and user configurations. It shows that AAA system (1) reduces the transmission delay, (2) increases the throughput, (3) improves the overall network utilization, (4) accommodates more users in the same network, and (5) provides load-balancing mechanism.

At the end, we conclude this paper with some possible future work. First, we note that the AAA system is yet to be put into practice. If the system is implemented on the access points/base stations, the network load can be easily controlled and the users will be better served within the wireless network. In addition, AAA system can cooperate with SNMP (Simple Network Management Protocol) to gather more precise information for reference when *CalBest* algorithm makes its decisions. Moreover, users may send a service agreement which indicates the maximum and minimum bandwidth needed for different classes of service to the SA for accurately reserve the resource. Also, SNMP here can be a great help for providing the necessary information to the reservation. Finally, since the AAA system is in server–client architecture, we expect to eliminate the centralized framework by introducing a routing protocol which is like RIP, OSPF, IGRP or IGRP to enable the communications between CAs. With the support of this protocol, the system can be further simplified.

References

Ada, I., Castellucia, C., 2001. Differentiation mechanisms for IEEE 802.11. In: Proceedings of IEEE Infocom'01, March, 2001.

Agere Systems Inc. NDIS 5 USB Driver for MS Windows XP, Winter, 2002.

Barry, M., Campbel, A.T., Veres, A., 2001. Distributed control algorithms for service differentiation in wireless packet networks. In: Proceedings of IEEE Infocom'01, April, 2001.

Bianchi, G., Fratta, L., Oliveri, M., 1996. Performance evaluation and enhancement of the CSMA/CA MAC Protocol for 802.11 Wireless LANs. In: Proceedings of the Seventh International Symposium on Personal, Indoor, and Mobile Communication (PIMC), 1996.

Boingo Software, version 1.1, 2002.

Chen, S., Nahrstedt, K., 1999. Distributed quality of service rouring in ad hoc networks. In: Proceedings of IEEE JSAC, August, 1999.

Cisco Aironet 350 Series (Aironet AIR-PCM 352 Wireless Lan Adapter), March, 2004.

Grow, B.P., Widjaja, I., Kim, J.G., Sakai, P., 1997. Investigation of the IEEE 802.11 medium access control (mac) sublayer functions. In: Proceedings of IEEE Infocom'97, April, 1997.

Hsiao, P., Hwang, A., Kung, H.T., Vlah, D., 2001. Load-balancing routing for wireless access networks. In: Proceedings of IEEE Infocom'01, April, 2001.

IEEE 802.11b/d3.0 Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, August, 1999.

IEEE 802.11 Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, August, 1999.

Kim, H., Jennifer, C. Hou., 2003. Improving protocol capacity with model-based frame scheduling in IEEE 802.11-operated WLANs. In: Proceedings of ACM MobiCom'03, September, 2003.

Kotz, D., Essien, K., 2002. Analysis of a campus-wide wireless network. In: Proceedings of ACM MobiCom'02, September, 2002.

Lindgren, Anders, Almquist, Andreas, Schelén, Olov, 2003. Quality of service schemes for IEEE 802. 11 wireless LANs: an evaluation. Mobile Networks and Applications 8 (3).

Mangold, S., Choi, S., May, Peter., Klein, Ole., Hiertz, Guido, Stibor, Lother., 2002. IEEE 802.11e wireless lan for quality of service. In: Proceedings of European Wireless'02, Italy, February, 2002.

Perkins, C. 1996. IP Mobility Support. RFC2002, IETF.

Shah, S.H., Nahrstedt, K., 2002. Predictive location-based QoS routing in mobile ad hoc networks. In: Proceedings of IEEE ICC'02, 2002.

Tang, D., Baker, M., 1999. Analysis of a metropolitan-area wireless network. In: Proceedings of ACM MobiCom'99, August, 1999.

Tang, D., Baker, M., 2000. Analysis of a local-area wireless network. In: Proceedings of ACM MobiCom'00, August, 2000.

The Network Simulator – ns-2, http://www.isi.edu/nsnam/ns/.

Tsai, J., Chen, T., Gerla, M., 1997. QoS routing performance in Multihop, multimedia, wireless networks. In: Proceedings of IEEE ICUPC, 1997.

Vaidya, N.H., Bahl, P., Gupta, S., 2000. Distributed fair scheduling in a wireless LAN. In: Proceedings of ACM MobiCom'00, August, 2000.

Veeraraghavan, T.M., Cocker, N., 2001. Support of voice services in IEEE 802.11 wireless LANs. In: Proceedings of IEEE Infocom'01, Polytechnic University, 2001.

Visser, M., ElZarki, M., 1995. Voice and data transmission over an 802.11 wireless network. In: Proceedings of IEEE Personal., Indoor and Mobile Radio Conference (PIMRC'95), September, 1995.

Zerfos, P., Zhong, G., Cheng, J., Luo, H., Lu, S., Li, J., 2003. DIRAC: a software-based wireless router system. In: Proceedings of ACM MobiCom'03, September, 2003.

Yaw-Chung Chen received his Ph.D. degree in Computer Science from Northwestern University, Evanston, Illinois in 1987. During 1987–1990, he worked at AT&T Bell Laboratories as a Member of Technical Staff. Since 1990, he has been with Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan. He is currently a professor. His research interests include multimedia communications, P2P systems, mobility in wireless networks, and Internet QOS. He is a senior member of IEEE and a member of ACM.

Wen-Yen Chen received his M.S. degree and B.S. degree in Computer Science from National Chiao Tung University in 2004 and 2001, respectively. He is currently a Ph.D. candidate in the Department of Computer Science at University of California, Santa Barbara. His recent research activities are in the areas of machine learning, data mining, parallel computing, and their applications to large-scale social-networks. Wen-Yen Chen is a recipient of the Research Creativity Award from National Science Council of ROC, and Regent Fellowship from the University of California.