

國立交通大學

網路工程研究所

碩士論文

無線網路電話之動態負載均衡方案

A Dynamic Load Balancing Scheme for VoIP over WLANs

研究生：徐誌謙

指導教授：曹孝櫟 教授

中華民國九十六年六月

無線網路電話之動態負載均衡方案
A Dynamic Load Balancing Scheme for VoIP over WLANs

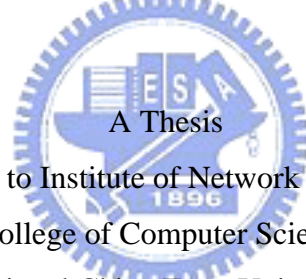
研究生：徐誌謙

Student：Chih-Chien Hsu

指導教授：曹孝櫟

Advisor：Shiao-Li Tsao

國立交通大學
網路工程研究所
碩士論文



A Thesis
Submitted to Institute of Network Engineering
College of Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in
Computer Science

June 2007

Hsinchu, Taiwan, Republic of China

中華民國九十六年六月

國立交通大學

博碩士論文全文電子檔著作權授權書

(提供授權人裝訂於紙本論文書名頁之次頁用)

本授權書所授權之學位論文，為本人於國立交通大學 網路工程 系所，
95 學年度第 二 學期取得碩士學位之論文。

論文題目：無線網路電話之動態負載均衡方案

指導教授：曹孝櫟 博士

同意 不同意

本人茲將本著作，以非專屬、無償授權國立交通大學與台灣聯合大學系統圖書館：基於推動讀者間「資源共享、互惠合作」之理念，與回饋社會與學術研究之目的，國立交通大學及台灣聯合大學系統圖書館得不限地域、時間與次數，以紙本、光碟或數位化等各種方法收錄、重製與利用；於著作權法合理使用範圍內，讀者得進行線上檢索、閱覽、下載或列印。

論文全文上載網路公開之範圍及時間：

本校及台灣聯合大學系統區域網路	<input checked="" type="checkbox"/> 中華民國 97 年 7 月 1 日公開
校外網際網路	<input checked="" type="checkbox"/> 中華民國 97 年 7 月 1 日公開

授權人：徐誌謙

親筆簽名：_____

中華民國 九十六 年 六 月 二十二 日

無線網路電話之動態負載均衡方案

學生：徐誌謙

指導教授：曹孝櫟

國立交通大學網路工程研究所碩士班

摘要

一般而言，無線網路擷取點的訊號涵蓋範圍通常會相互覆蓋，因此一無線網路用戶可能會偵測到多個無線網路擷取點並且選擇其中一個來連線以進入無線網路熱點。實驗結果顯示一個無線網路用戶通常會和訊號最強的擷取點聯繫並向其要求頻寬以便建立網路連線，然而，這種以用戶為中心的聯繫以及頻寬要求策略可能會導致擷取點間的負載不均，以致無線網路擷取點中擷取點的頻寬資源效用無法完全發揮。對於商業上部建無線網路電話系統這類必須在保證通話品質下最大化共同通話數目之無線網路電話系統而言，此一負載不均的問題將會是一個重要的議題。

本論文針對無線網路電話系統提出了一個之動態負載平衡方案，一開始網路協助的聯繫機制會建議一個用戶向負載最輕的擷取點來要求建立無線網路電話之通話。假使此用戶所有偵測到的擷取點皆負荷過載時，則本論文所提出的負載平衡方案將會啟動來進一步重新規劃正在通話的無線網路電話用戶和擷取點間的連線情形以便能騰出足夠的資源來提供新的通話要求。模擬結果說明了藉著本論文所提出的方案，無線網路電話系

統通話服務要求被拒絕建立通話的機率將可大幅度地降低。



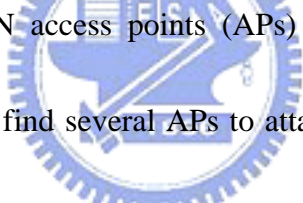
A Dynamic Load Balancing Scheme for VoIP over WLANs

Student: Chih-Chien Hsu

Advisor: Shiao-Li Tsao

Institute of Network Engineering
College of Computer Science
National Chiao Tung University

Abstract



Coverage areas of WLAN access points (APs) are usually overlapped so a WLAN station (STA) might be able to find several APs to attach in a WLAN hotspot. Experimental results indicate that a WLAN STA normally associates with an AP with the maximal signal strength and requests the bandwidth of the AP for establishing network connections. However, this kind of STA-centric association and bandwidth request policy may introduce unbalance loads of APs, and the bandwidths of APs in a WLAN hotspot cannot be fully utilized. This unbalance load problem is a critical issue for the commercial deployment of voice over IP (VoIP) over WLAN (VoWLAN) service which has to maximize the number of concurrent VoWLAN sessions with Quality of Service (QoS) guarantees. In this paper, a novel dynamic load balancing scheme is proposed for a VoWLAN system. The network-assisted association policy first advises an STA to request a VoWLAN session through an AP with the minimal

load. In case of the APs which the STA can attach are all overloaded, the proposed load balancing scheme further rearranges the serving VoWLAN STAs between APs in order to spare enough resources for accommodating that new request. Simulation results demonstrate that the reject rate of service requests for a VoWLAN system can be considerably reduced by employing the proposed scheme.



誌謝

本篇論文能順利完稿，要感謝的人很多，首先我要感謝的是我的指導教授—曹孝櫟老師。不論是在研究或待人處事上，老師都對我有深遠的影響，因著老師的悉心教誨與照顧提攜，我才能有充實且精彩的研究所生活，如果說我在這兩年有些許的成績表現，這些都是老師所賦予我的，在此獻上我由衷的謝意。接著，我要感謝口試委員曾建超教授、鄭瑞光教授以及徐玉青博士的指正與建議，讓我獲益匪淺。

同時，我要感謝交大資工系 BRASS 實驗室的成員。謝謝建明學長、宥霖學長以及一正學長在論文的撰寫過程中，給予我極大的幫助，使這篇論文能更加完整。謝謝政龍、中暉、名杰、雅筑和彥筑，我不會忘記這段一起相互討論，相互勉勵的日子。謝謝建臻、家祥以及世永學弟協助處理許多事情，讓我可以專心致力於研究。謝謝你們在這段期間對我的照顧。

最後，我要感謝我的家人。謝謝父母養育並栽培我，沒有你們，就沒有今日的我。謝謝誌宏兄長從小就為我樹立了良好的典範，引領我在求學過程中能不迷失自己的方向。謝謝以琳這一路來的支持與鼓勵，因為妳的陪伴，讓我在研究過程中，無論是遇到挫折而失落或是突破瓶頸而欣喜若狂時，總是能夠有個人和我分享。

感謝所有幫助與關心我的人，僅以本篇論文與大家分享。

目錄

摘要	i
Abstract	iii
誌謝	v
目錄	vi
圖目錄	vii
符號說明	viii
1. Introduction	1
2. Dynamic Load Balancing Scheme.....	4
<i>2.1. Dynamic load balancing example</i>	<i>4</i>
<i>2.2. Modeling the relationships between APs and STAs</i>	<i>6</i>
<i>2.3. Dynamic load adjustment</i>	<i>8</i>
3. Simulation Results and Analysis	15
4. Conclusions	23
References.....	24



圖目錄

Figure 1. Dynamic load balancing examples.....	6
Figure 2. The resource-allocation graph of Figure. 1(a).....	9
Figure 3. The resource-allocation graph of Figure 2 after performing load adjustments	11
Figure 4. The proposed admission control and the dynamic load balancing scheme	12
Figure 5. The admission control and proposed dynamic load balancing scheme	13
Figure. 6. Reject rate of service requests which are rejected under different load conditions and a WLAN hotspot with $D = 1.5$	19
Figure. 7. Reject rate of service requests which are rejected under different load conditions and a WLAN hotspot with $D = 3.0$	20
Figure. 8. Reject rate of service requests which are rejected under different load conditions and a WLAN hotspot with $D = 6.0$	21
Figure. 9. The overheads by employing the proposed scheme for a WLAN hotspot with $D = 1.5, 3.0$ and 6.0	22



符號說明

- A_i : i^{th} AP in a WLAN hotspot
 S_j : j^{th} STA in a WLAN hotspot
 C_i : the current resource utilization of i^{th} AP
 $R_{i,j}$: the speed which j^{th} STA associates i^{th} AP at
 r : the bandwidth which an STA needs for a VoWLAN session
 $n_{i,j}$: the coverage relationships between APs and STAs
 $m_{i,j}$: the serving relationships between APs and STAs



1. Introduction

The technology development and network deployment of WLANs have grown rapidly in recent several years, and WLAN has become one of the most popular access technologies for mobile Internet services. Among all mobile Internet services and applications, voice over IP (VoIP) over WLAN (VoWLAN) has attracted considerable interest from both academia and industry and is regarded as one of the killer applications for both public and enterprise WLANs [1]. However, VoWLAN applications generate a large amount of small voice packets which degrade the WLAN utilization due to the nature of WLAN medium access control (MAC) mechanism [7], and the service capacity of a WLAN access point (AP) for VoWLAN services, i.e. the number of concurrent VoIP sessions that a WLAN AP can support, is very limited [2]. To increase the number of VoWLAN sessions that an AP can serve and to maximize the resource utilizations of APs in a WLAN hotspot are both challenging issues for the commercial deployment of VoWLAN services.

Previous studies have worked on improving the utilization of a single AP for VoWLAN services [3], and have proposed several radio resource management schemes for WLANs [4, 5, 6]. Considering a WLAN hotspot with multiple APs whose coverage areas are overlapped, the AP association policies for STAs become important since a WLAN STA might be able to find several APs to attach. Conventional AP association policies are usually implemented on STAs

and suggest STAs to select and request bandwidth of an AP with the maximal signal strength. However, this STA-centric network association and service request policy may introduce unbalance loads on APs, and the utilizations of APs in a hotspot cannot be maximized [17, 18, 19]. To improve the overall utilization of APs in a WLAN hotspot, research suggests network-assisted approaches to manage WLAN resources. A number of studies suggest APs to broadcast the load information of an AP in Beacon or Probe Response frames and advice STAs to attach to the AP with the minimal load [11, 12]. Besides assisting newly arrival STAs to associate with the AP with the minimal load, research studies in [10, 13, 14, 15, 16] further request the STAs which already attach the APs but situate in poor wireless link conditions to roam from the serving AP to the neighbor APs. The policies which dynamically assign the STAs to associate with APs with better link qualities and/or the minimal load benefit the overall utilization of a WLAN hotspot. Unfortunately, some of above studies assume STAs requesting best effort services and contending WLAN channels with peer STAs, but they did not consider the resource management for QoS sessions such as VoWLAN service. These network-assisted mechanisms are not suitable to be directly applied to a VoWLAN system. Although some other approaches propose the resource management schemes for QoS services, these schemes only handle local load balancing for APs. The global optimization issue and the load rearrangement among all APs in a WLAN hotspot have not yet considered.

In this work, a novel dynamic load balancing scheme is proposed for a VoWLAN system

which offers guaranteed QoS services. The network-assisted mechanism first advises an STA to associate with an AP with the maximal available resources. In case of the APs that the STA can attach are all overloaded, the proposed load balancing scheme is activated to adjust the loads between APs in order to accommodate that new VoWLAN request. The rest of the paper is organized as follows. The concept and procedures of the proposed dynamic load balancing scheme for a VoWLAN system are presented in Section 2. Simulation results are discussed in Section 3, and finally Section 4 concludes this work.

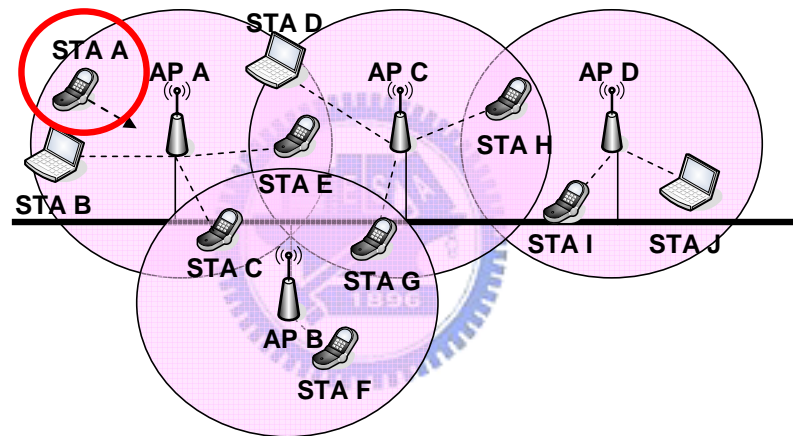


2. Dynamic Load Balancing Scheme

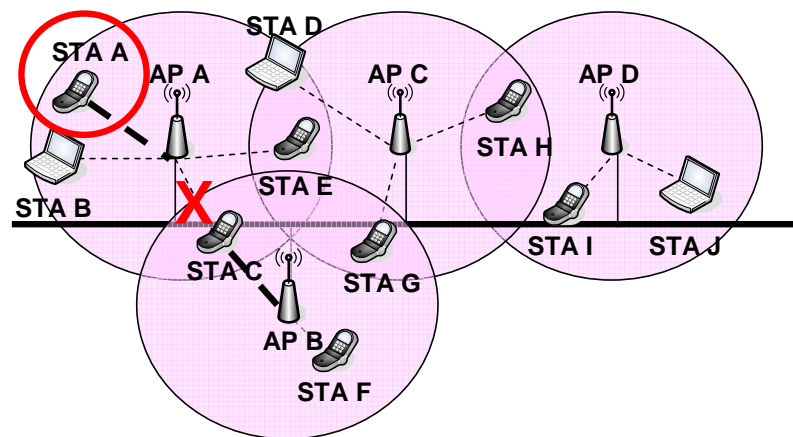
2.1. Dynamic load balancing example

An STA, say STA A, is able to associate with a WLAN AP, say AP A, only. While AP A is overloaded, the STA A cannot obtain enough resources from AP A and the service request from STA A is thus rejected by AP A. Considering AP A is currently serving another STA, say STA B, and STA B can find and associate with another AP, say AP B, which is under-loaded. STA B can change its serving AP from AP A to AP B, and then the resources occupied by STA B on AP A can be released. Therefore, the resources on AP A now become available to serve the new STA A. The load adjustment procedure can be done between two APs, and it can be also applied to a chain of multiple APs. Figure 1 (a) shows an example where circles represent the coverage areas of APs and the adjoined APs occupy different WLAN channels. In this example, each AP is assumed to support at most three VoIP sessions. STAs B, C, D, E, F, G, H, I and J associate with APs A, A, C, A, B, C, C, D and D respectively. While STA A that can only associate with AP A attaches to the network and requests VoWLAN services, AP A which is overloaded cannot provide the service. A dynamic load balancing scheme is thus applied to the situation. The scheme changes the serving AP of STA C from AP A to AP B, and then the resources on AP A become available to be allocated to STA A. Therefore, the service request from STA A can be accepted by AP A. Figure 1 (b) illustrates the example shown in Figure 1

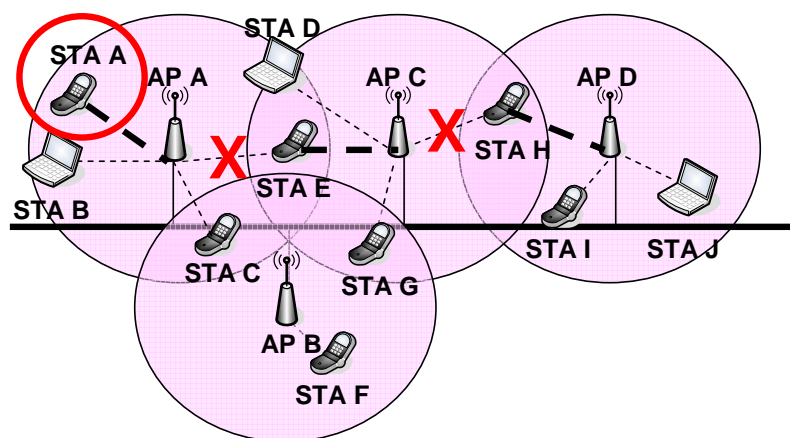
(a) after applying the proposed dynamic load balancing scheme. Figure 1 (c) shows another example that a dynamic load adjustment can be applied to a chain of multiple APs. While STA A requests a VoIP session to AP A, and AP A is overloaded. STA H can change its AP from AP C which is also overloaded to AP D which is under-loaded. AP C has available resources to serve new requests, and then STA E can change its serving AP from AP A to AP C. Therefore, the resources on AP A become available to be allocated to STA A.



(a) Before load adjustment, STA A cannot obtain the resources



(b) After migrating STA C from AP A to AP B, STA A can obtain the resources



(c) After migrating STA H from AP C to AP D, STA E from AP A to AP C, STA A can obtain the resources.

Figure 1. Dynamic load balancing examples

2.2. Modeling the relationships between APs and STAs

Before the dynamic load balancing scheme is presented, the relationships between STAs and APs in a WLAN hotspot are first modeled. A WLAN hotspot totally contains N WLAN APs, and all APs in the hotspot are assumed identical. The current resource utilization of i^{th} AP, say A_i , is denoted as C_i which is a value between 0 and 1. C_i is defined as the percentage of transmission time occupied by all serving STAs over the total operation time of A_i . Therefore, $C_i=1$ implies all resources on A_i are occupied by the STAs and there is no resource available to serve any new service. The j^{th} STA, denoted as S_j , associates with a WLAN AP, say A_i , at $R_{i,j}$ speed in Kbps. For example, the IEEE 802.11b offers 1Mbps, 2Mbps, 5 Mbps, and 11Mbps speeds for STAs, and the association speed between an AP and an STA depends

on the distance and channel quality between them. Assume that an STA requests a VoWLAN session at r Kbps, and then the AP has to allocate $r/R_{i,j}$ resources for the VoWLAN session if A_i admits S_j . The above equations use very simple models to evaluate the resource consumed by an STA which associates with an AP at $R_{i,j}$ Kbps and requests r Kbps for its VoIP session. To precisely calculate the resources occupied by an STA for the AP, studies have proposed a number of models. For example, IEEE 802.11e [9] suggests a resource consumption model while performing the call admission control for hybrid coordination function (HCF) controlled channel access (HCCA) mechanism. In this mechanism, an STA, say S_j , must send a request message (ADDTS) to its serving AP, say A_i , for establishing a VoIP session. The request message contains a traffic specification (TSPEC) which describes the traffic characteristics and the QoS requirements of the VoIP session. After the serving AP receives the request, it evaluates its available resources and decides to accept this request or not. If the request is accepted, the AP replies the STA with transmission opportunity (TXOP) duration among a scheduled service interval (SI). In other words, if the new request is admitted, the AP has to allocate $\frac{TXOP}{SI}$ resources to that new request. Different from the simple resource model described above, IEEE 802.11e [9] provides more accurate resource models to calculate TXOP. More precise and accurate resource models and admission control schemes can be incorporated with and applied to the proposal dynamic load balance approach. Without loss of generality, the simple model is used to present the basic idea behind the proposed load

adjustment scheme.

Besides the resource models, the relationships between APs and STAs are also defined.

While an STA, say S_j , performs WLAN channel scan and finds an AP, say A_i , S_j then inserts

A_i into the candidate list. Here, $n_{i,j}$ defines the coverage relationships between APs and STAs

as:

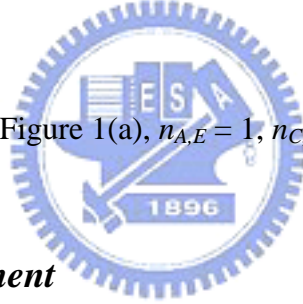
$$n_{i,j} = \begin{cases} 1, & \text{if } A_i \text{ is in } S_j \text{' candidate list.} \\ 0, & \text{otherwise.} \end{cases}$$

After scan procedures, S_j decides to associate with A_i , and requests a VoWLAN service,

$m_{i,j}$ defines the serving relationships between APs and STAs as:

$$m_{i,j} = \begin{cases} 1, & n_{i,j} = 1 \text{ and } A_i \text{ is serving } S_j. \\ 0, & \text{otherwise.} \end{cases}$$

For example, for STA E in Figure 1(a), $n_{A,E} = 1$, $n_{C,E} = 1$ and $m_{A,E} = 1$, $m_{C,E} = 0$.



2.3. Dynamic load adjustment

To achieve a better network utilization, a network-assisted policy assigns the AP with the minimal load to serve a new VoWLAN STA. That is, a new VoWLAN STA, say S_j , is asked to associate with A_i that is in S_j 's candidate list, i.e. $n_{i,j}=1$, and A_i has the maximal available resources after serving S_j , i.e. A_i with the minimal $C_i+r/R_{i,j}$. If the serving AP of a VoWLAN STA does not have enough resources, the traditional network-assisted approaches reject this VoWLAN request. In our design, the second step procedure, i.e. the dynamic load balancing procedure, is activated to adjust loads between APs to accommodate that new request. A direct graph is newly proposed in this paper to represent the current loads and the relationships

between APs and STAs. The directed graph, called resource-allocation graph G , illustrates the resource-allocated status between APs and STAs. Vertices V in graph could be STAs or APs.

An edge E denotes the relationship between vertices. An edge only appears between one AP and one STA, but does not exist between two APs or two STAs. An edge from A_i to S_j denotes as (A_i, S_j) means AP A_i is serving STA S_j , called an assignment edge. That is $n_{i,j}=1$ and $m_{i,j}=1$.

If there is an edge from S_j to A_i , denoted as (S_j, A_i) , implies A_i is in S_j 's candidate list but A_i is not serving S_j , called a claim edge. In other words, $n_{i,j}=1$ and $m_{i,j}=0$. Figure 2 shows an

example of the resource-allocation graph for Figure 1 (a). The relationship between APs and

STAs can be easily obtained from the resource-allocation graph, and the dynamic load balancing scheme can use the graph to determine the load adjustments between APs.

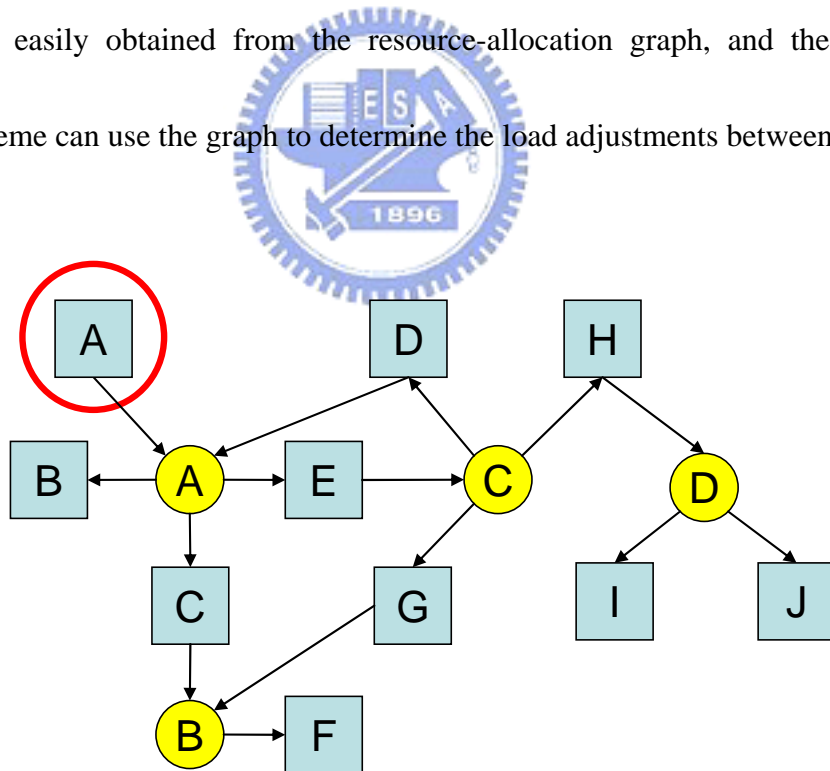


Figure 2. The resource-allocation graph of Figure. 1(a)

It can be seen from Figure 2 that STA A can be only served by AP A but unfortunately AP A is overloaded. The next step of the load balancing procedure is to find a directed path

without an STA visited twice in the resource-allocation graph from STA A to any other AP with available resources. A feasible path P in G for S_j is defined as $\{(S_j, A_i), (A_i, S_j'), (S_j', A_i'), \dots, (S_j^{(n)}, A_i^{(n)})\}$. In this path, A_i to $A_i^{(n-1)}$ could be overloaded and only $A_i^{(n)}$ must be under-loaded and can serve $S_j^{(n)}$. A path represents a list of load adjustment operations. For example, S_j' can change its current serving AP from A_i to A_i' , and then A_i has available resources to serve S_j . Before the migration of S_j' , S_j'' can change its serving AP from A_i' to A_i'' . Since $A_i^{(n)}$ is under-loaded, STAs S_j' to $S_j^{(n)}$ can perform migrations for the old serving APs to the new serving APs. Therefore, the new STA S_j can be admitted and served by A_i . A VoWLAN request might have multiple feasible paths. For example, the case shown in Figure 3 has at least two feasible paths, i.e. $\{(STA A, AP A), (AP A, STA C), (STA C, AP B)\}$ and $\{(STA A, AP A), (AP A, STA E), (STA E, AP C), (AP C, STA H), (STA H, AP D)\}$. If more than one path is found, the shortest path which implies the minimal migration overheads is selected. Once the path is decided, the direction of edges of the path should be reversed. That is, assignment edges become claim edges and claim edges become assignment edges. Figure 3 illustrates the resource-allocation graph of Fig. 1(c) after the load adjustment is performed. While STA A cannot get the resources from AP A, the dynamic load balancing mechanism is to find a load adjustment path P from STA A to AP D: $\{(STA A, AP A), (AP A, STA E), (STA E, AP C), (AP C, STA H), (STA H, AP D)\}$. Then, the directions of the edges in the path should be reversed. Figure 4 illustrates the flow chart for the admission control and the

procedures for the proposed dynamic load balancing scheme. The blocks with different colors illustrate the procedures that the STA and APs should perform. Figure 5 shows the algorithm for the admission control and dynamic load balance scheme.

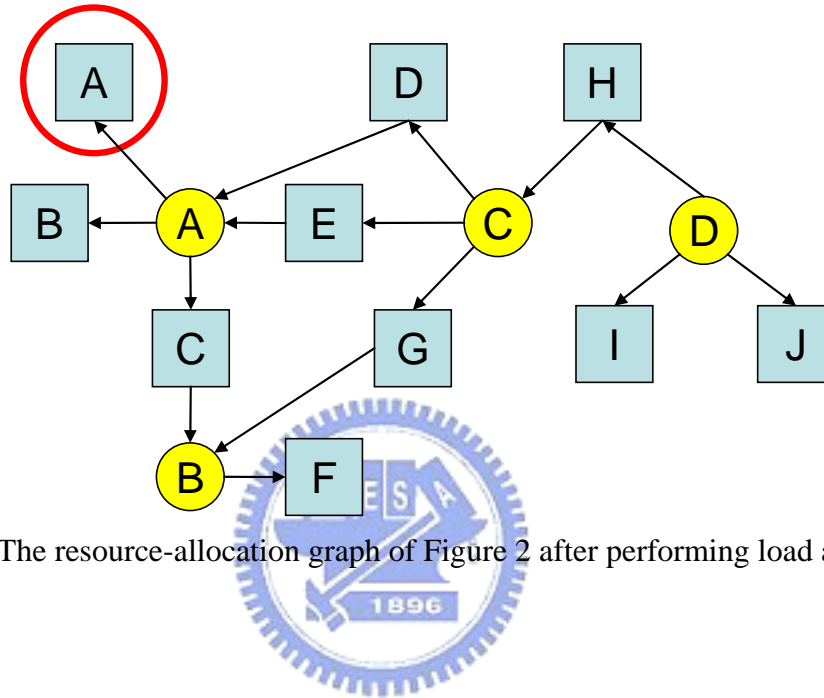


Figure 3. The resource-allocation graph of Figure 2 after performing load adjustments

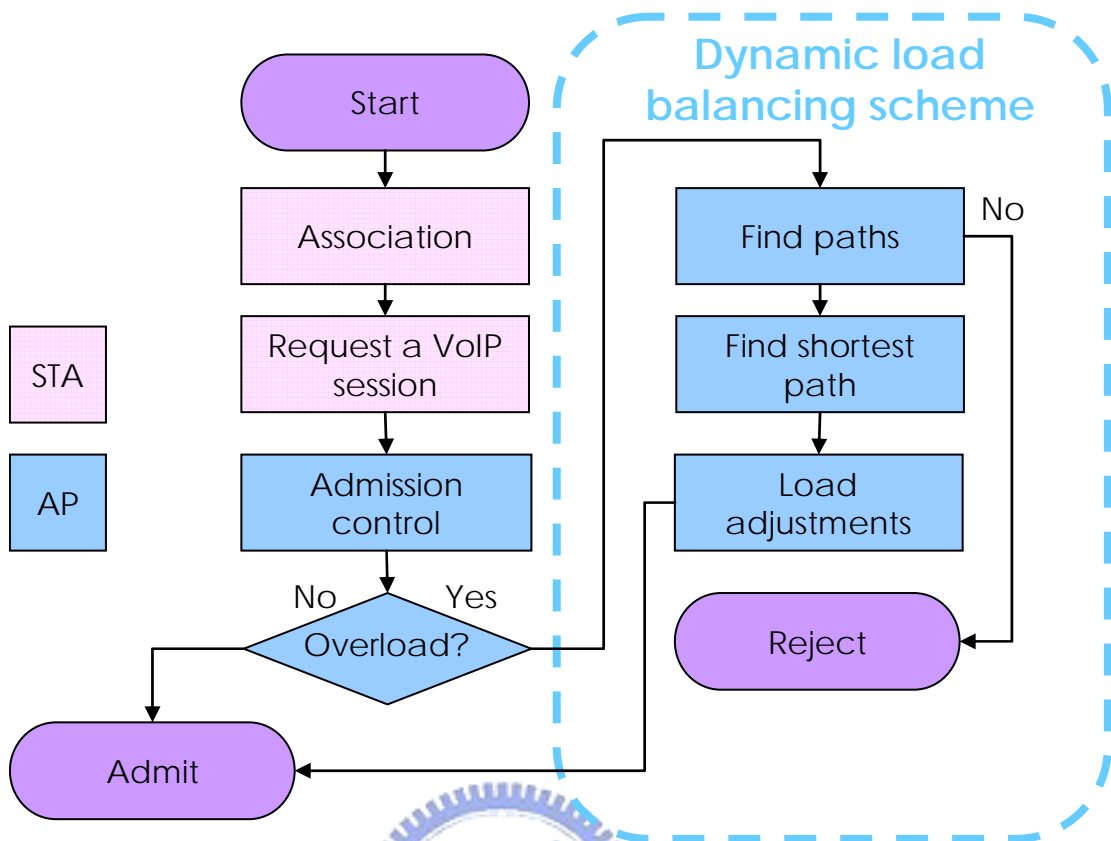


Figure 4. The proposed admission control and the dynamic load balancing scheme



```

CP: set of feasible paths
P: a feasible path
SP: the shortest feasible path
V: set of vertices have been visited

Admission Control ( $S_j$ )
{ if (Overload( $S_j$ ) = false)
  { if ((CP=Find Path( $S_j, R_{i,j}$ )) = true)
    { SP=Find Shortest Path(CP); //return the shortest path
      Load Adjustment(SP); //reverse the directions of the edges
in the path.
    }
    else //can not find any feasible path
      Reject the request;
  }
  else //APs with available resources found
    Accept the request;
}

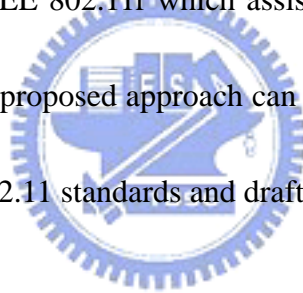
Overload( $S_j$ )//checks if all APs that STA can find are overloaded
{ min_C=1;
  for i where  $n_{i,j}=1$  //calculate resources
    { if  $C_i+r/R_{i,j} \leq 1$  and  $C_i+r/R_{i,j} < \text{min\_C}$ 
      min_C= $C_i+r/R_{i,j}$  and min_A=i
    }
  if min_C=1
    return false //all APs the STA can hear are overloaded
  else
    return min_A //return AP with the minimal load
}

Find Path( $S_j, R_{i,j}$ )// depth first traversal
{ Add  $S_j$  to V;
  for i where  $n_{i,j}=1$  and  $m_{i,j}=0$  AND  $A_i \notin V$ 
    { Add  $A_i$  to V;
      Add ( $S_j, A_i$ ) to P;
      //Search for STA  $S_k$  from the graph where  $m_{i,k}=1$ 
      if ( $C_i+r/R_{i,j} \leq 1$ )
        Add P to CP; //path is found
      else if(  $S_k \notin V$  and  $C_i+r/R_{i,j}-r/R_{i,k} \leq 1$ ){
        { Add ( $A_i, S_k$ ) to P;
          Find Path( $S_k, R_{i,k}$ );
          Delete ( $A_i, S_k$ ) from P;
        }
        Delete  $A_i$  from V;
        Delete ( $S_j, A_i$ ) from P;
      }
    }
  Delete  $S_j$  from V;
}

```

Figure 5. The admission control and proposed dynamic load balancing scheme

To implement the proposed dynamic load balancing scheme, one possible approach is to setup a centralized server for admitting service requests and initiating the load adjustments between APs. A distributed approach which exchanges the load conditions of APs and performs the load adjustments is also possible. These load information of APs and the control messages for performing load adjustments between APs are exchanged over the backbone network. To implement the STA migration between APs, IEEE 802.11f [20] which can transfer the context of an STA between APs, IEEE 802.11k [8] which can advise an STA to attach to a specific AP, and IEEE 802.11r which assists APs and STAs to perform seamless handovers can be utilized. The proposed approach can be implemented in WLAN hotspots by integrating the current IEEE 802.11 standards and draft standards.



3. Simulation Results and Analysis

Simulations are conducted to evaluate the performance by applying the conventional STA-centric approach, the network-assisted approach and the proposed dynamic load balancing scheme. For the conventional STA-centric approach, an STA always sends its request to the AP with the maximal signal strength. If the AP cannot accept this request, the request is rejected. For the network-assisted approach, the AP actively broadcasts the loads of APs, and advises the STA to request its QoS session to the APs with the minimal load. If all APs that the STA can attach are fully occupied, the service request is rejected. For the proposed dynamic load balancing scheme, an STA first sends its request to the AP with the minimal load. If the request is rejected, the dynamic load balancing scheme is activated to find a feasible load adjustment in order to accommodate the new request. If there is no load adjustment can be performed, the service request is rejected. The reject rate of the new service requests and the overhead which is introduced by employing the proposed method are both investigated. The reject rate of the new service requests is the percentage of new requests which are rejected by APs. The overhead here is defined as the numbers of STAs which are forced to change their serving APs in order to accommodate new requests.

In a simulation, a deployment scenario of a WLAN hotspot is first generated. A deployment scenario means a particular number of WLAN APs which are randomly deployed

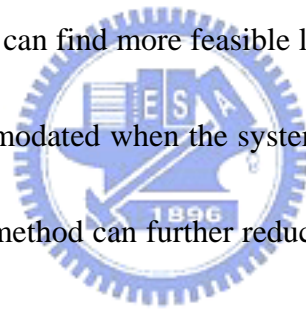
in a fixed-size hotspot. To simplify the simulations, an AP only offers one association speed, i.e. 11Mbps, within the coverage area of an AP, which is a 30-meter-radius range. A WLAN hotspot is a 300 meters by 300 meters square area. In our simulations, three different network densities D which are $D=1.5$, $D=3.0$ and $D=6.0$ of WLAN hotspots are considered. The network density D here is defined as the average number of APs that an STA can detect at any location of a WLAN hotspot. For example, IEEE 802.11b/g has three non-overlapped channels, and operators may install IEEE 802.11b/g APs in a hotspot where STAs can hear APs in the three non-overlapped channels, i.e., Channel 1, Channel 6 and Channel 11. In such a deployment, the network density D could be approximately three. For IEEE 802.11a, there are twelve non-overlapped channels, the network density D could be twelve. If a WLAN hotspot installs both IEEE 802.11b/g and IEEE 802.11a APs, the network density D could be up to 15.

After a deployment scenario of a WLAN hotspot is settled, STAs that appear at random locations within a WLAN hotspot and send service requests to APs for establishing QoS connections are generated. An STA requests only one G.711 VoIP session which consumes 80Kbps downlink and 80Kbps downlink bandwidth of a WLAN AP. The length of a VoIP session is randomly generated between one to 30 minutes, and the call occupies the resources for the entire VoIP session. The arrival of the service requests is assumed a Poisson process. Different arrival rates of the service requests are generated in order to simulate different loads

of APs in the hotspot. The QoS of service requests from STAs are all identical. Therefore, each AP can support up to eight VoIP sessions concurrently. All simulations are based on the average results which are collected from a total of 100 randomly deployed scenarios of WLAN hotspots.

First, the percentage of service requests which are rejected by APs are evaluated under different loads of APs in a WLAN hotspot and different network densities. Figure 6 shows the reject rate of the service requests under different loads of WLAN APs and $D=1.5$. It can be learned from the figure while the loads of APs are low, e.g. less than 30%, the reject rates of a system by applying the three approaches are similar and all service requests can be accepted. When the system load increases, the network-assisted approach and the proposed approach can reduce the reject rate of the service requests. Although, the network-assisted approach and the proposed approach both reduce the reject rate of the service requests, the improvements are marginal. That is because in these deployment scenarios with $D=1.5$, an STA can detect only one or two APs. It is very difficult for STAs to find many alternative APs to attach. The reject rate of the service requests cannot be improved too much by applying the network-assisted approach. Moreover, the proposed approach has to find STAs which can attach to more than one AP and then the algorithm can find load rearrangements between APs. Without enough network densities, the benefit that the proposed approach can gain is very limited. Thus, we change the network density from $D=1.5$ to $D=3.0$. Figure 7 shows the

simulation results. It can be seen from the figure that both the network-assisted approach and the proposed approach reduce the reject rates of the service requests. The performance is significantly improved while the system is heavily loaded, i.e. 60% to 90%. For the network assisted approach, to select an AP with the minimal load can reduce the service request rate. The proposed scheme further achieves lower service reject rates than that of the network-assisted approach by rearranging the loads between APs. Simulation results show that 10% improvement compared to the network-assisted approach can be achieved by employing the proposed scheme under the system load is 80%. For STAs can find more APs to attach, the proposed scheme can find more feasible load adjustment paths so that more new service requests can be accommodated when the system load is heavy. If the network density increases to six, the proposed method can further reduce the reject rate of the service requests than the network-assisted approach by 30% under the system load is 90%. Figure 8 illustrates the simulation results. We can conclude that for hotspots with high network densities and are heavily loaded, i.e. 60% to 90%, the proposed mechanism significantly minimizes the reject rate of the service requests.



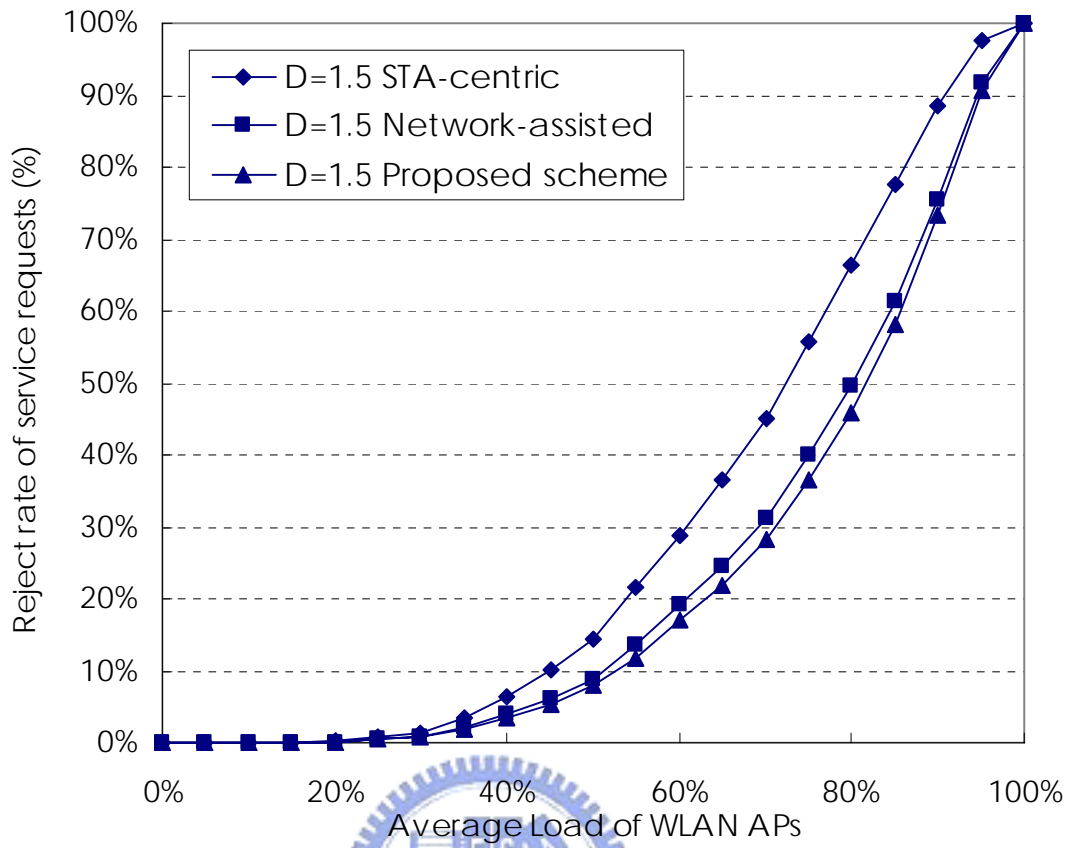


Figure. 6. Reject rate of service requests which are rejected under different load conditions and a WLAN hotspot with $D = 1.5$

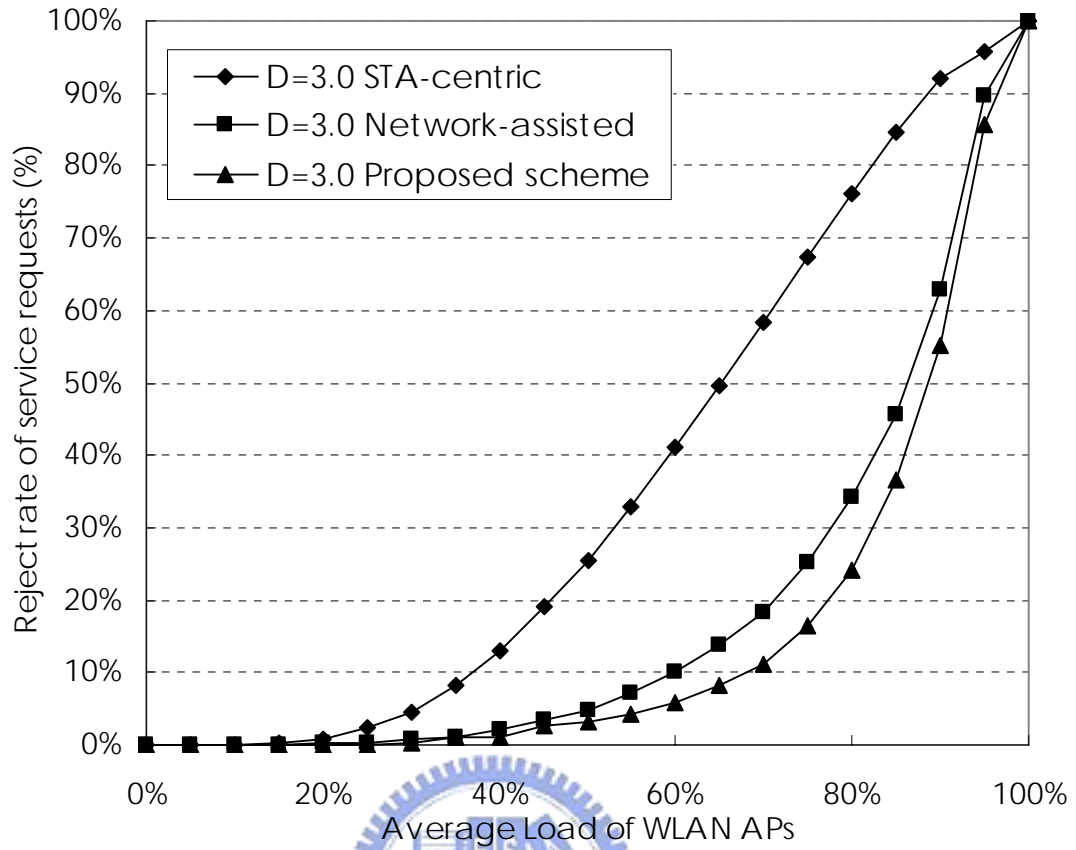


Figure. 7. Reject rate of service requests which are rejected under different load conditions and a WLAN hotspot with $D = 3.0$

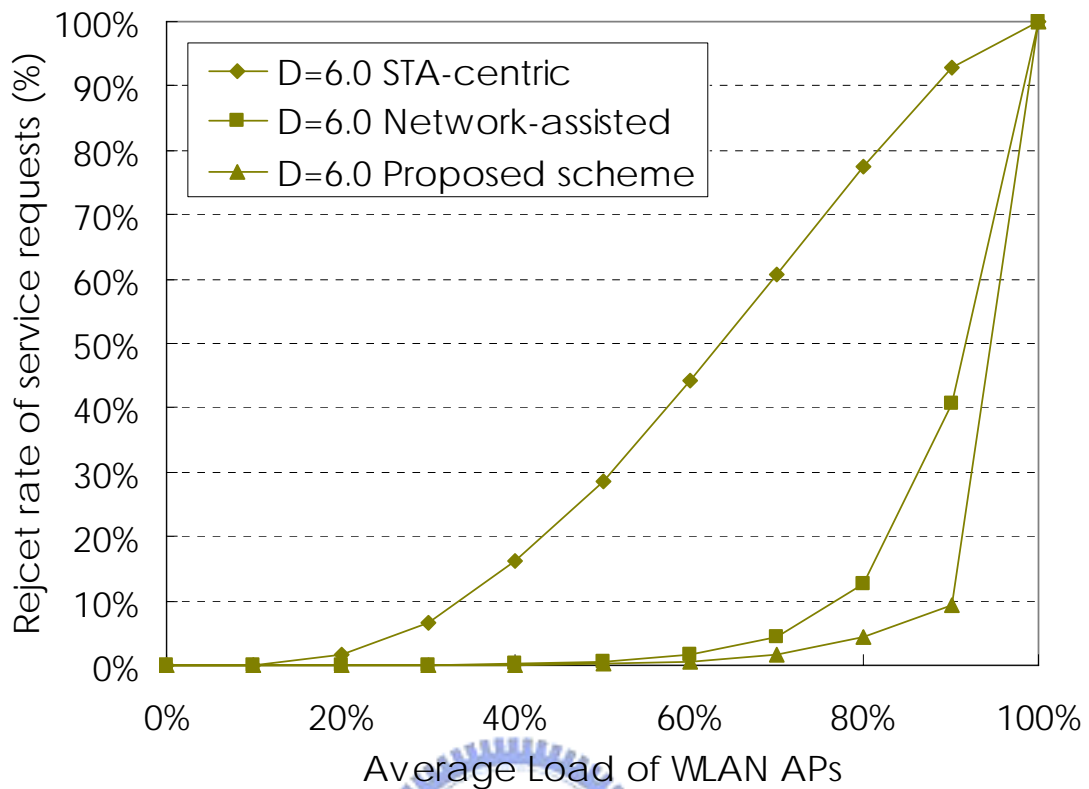


Figure. 8. Reject rate of service requests which are rejected under different load conditions and a WLAN hotspot with $D = 6.0$

Then, the overhead by employing the proposed scheme is evaluated. Figure 9 demonstrates the overheads of the proposed scheme in terms of roamed STAs under different network densities. The number of roamed STAs means that the number of existing STAs which have to roam from one AP to another AP in order to accommodate the new requests. It can be seen from Figure 9 that the numbers of roamed STAs increase while the network density and workload increase. This is because the proposed method is especially useful while network load and density are both high. For the system is almost fully loaded, i.e. more than 90%, the proposed method can not improve anymore since all requests are rejected. Simulation results demonstrate when the system load is 60% to 90%, only 1.5 to 2.5 STAs are

influenced for accommodating a new request under $D=3.0$, and only 2.5 to 4 STAs have to roam from their current APs to other APs under $D=6.0$.

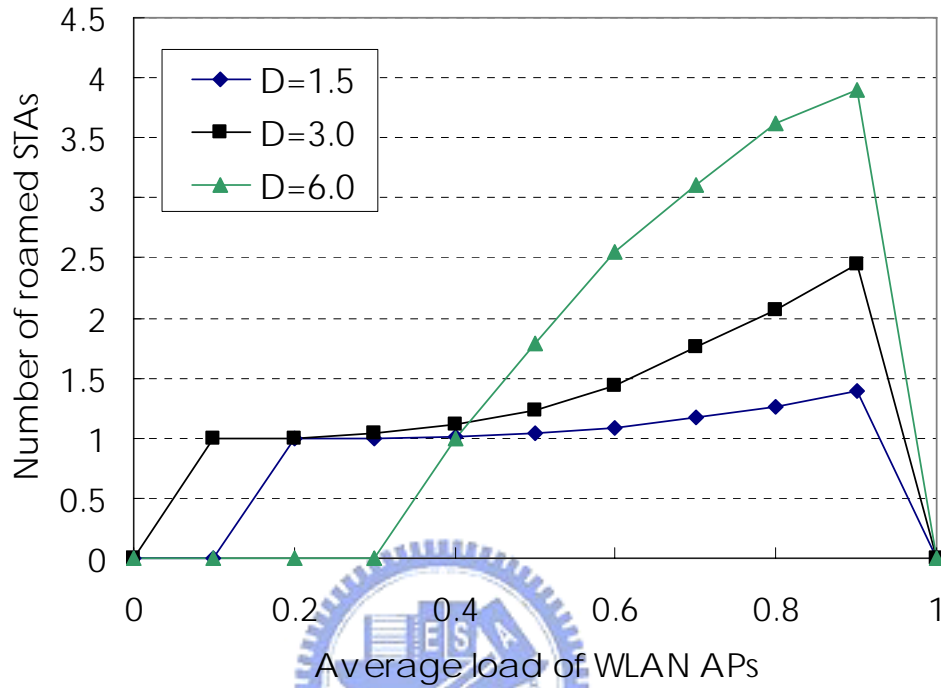
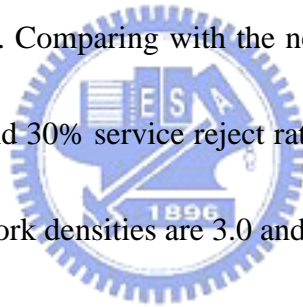


Figure. 9. The overheads by employing the proposed scheme for a WLAN hotspot with $D = 1.5, 3.0$ and 6.0

4. Conclusions

In this study, a novel dynamic load balancing scheme was proposed for a VoWLAN system. The network-assisted policy balances the loads of APs, and the dynamic load balancing scheme further reduces service reject rates when the system is heavily loaded. The proposed scheme rearranges the serving STAs between APs in a WLAN hotspot in order to accommodate the service requests. Simulation results demonstrate that the proposed approach can reduce the service reject rate of the conventional STA-centric approach by 20% to 54% under different load conditions. Comparing with the network-assisted approach, the proposal scheme further reduces 10% and 30% service reject rates while the loads of a WLAN hotspot are 80% and 90% and the network densities are 3.0 and 6.0, respectively.



References

- [1] D. Lenton, "Speaking of Wi-Fi," IEE Review, Vol. 49, Iss. 7, pp. 44 - 47, July 2003.
- [2] Moncef Elaoud et al., "Experimental VoIP Capacity Measurements for 802.11b WLANs," in Proceeding of IEEE Consumer Communications and Networking Conference (CCNC), 2005.
- [3] Wei Wang, Soung Chang Liew, and Li, V.O.K., "Solutions to Performance Problems in VoIP over a 802.11 Wireless LAN," IEEE Transactions on Vehicular Technology, Vol. 54, Iss. 1, pp. 366-384, Jan. 2005.
- [4] Bo Li, Li Yin, K. Y. Michael Wong, and Si Wu, "An Efficient and Adaptive Bandwidth Allocation Scheme for Mobile Wireless Networks Using an On-Line Local Estimation Technique," Wireless Networks, Vol. 7, Iss. 2, March 2001.
- [5] Mark Davis, "Radio Resource Management: A Wireless Traffic Probe for Radio Re-source Management and QoS Provisioning in IEEE 802.11 WLANs," in Proceeding of the 7th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems, October 2004.
- [6] Hakima Chaouchi and Anelise Munaretto "Adaptive QoS Management for IEEE 802.11 Future Wireless ISPs," Wireless Networks, Vol. 10, Iss. 4, July 2004.
- [7] IEEE 802.11 Standard, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification," 1999.
- [8] IEEE 802.11k Draft Standard, "Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 9: Radio Resource Measurement," 2005.
- [9] IEEE 802.11e-2005, "Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements," 2005.
- [10] Yigal Bejerano, Seung-Jae Han, and Li (Erran) Li, "Fairness and Load Balancing in Wireless LANs Using Association Control," in Proceeding of ACM MobiCom'04, 2004.
- [11] Proxim Wireless Networks. ORINOCO AP-600 Data Sheet, 2004.

- [12] Cisco Systems Inc. Data Sheet for Cisco Aironet 1200 Series, 2004.
- [13] I. Papanikos and M. Logothetis, "A Study on Dynamic Load Balance for IEEE 802.11b Wireless LAN," in Proceeding of International Conference on Advances in Communication and Control, 2001.
- [14] A. Balachandran, P. Bahl, and G. M. Voelker, "Hot-Spot Congestion Relief and Service Guarantees in Public-Area Wireless Networks," SIGCOMM Computer Communication Review, Vol. 32, No. 1, pp. 59-59, 2002.
- [15] T.-C. Tsai and C.-F. Lien, "IEEE 802.11 Hot Spot Load Balance and QoS-Maintained Seamless Roaming," in Proceeding of National Computer Symposium (NCS), 2003.
- [16] H Velayos, V. Aleo, and G. Karlsson, "Load Balancing in Overlapping Wireless LAN Cells," in Proceeding of IEEE ICC, France, June 2004.
- [17] A. Balachandran, G. M. Voelker, P. Bahl, and P. V. Rangan., "Characterizing User Behavior and Network Performance in a Public Wireless LAN," in Proceeding of ACM SIGMETRICS, pp. 195-205, 2002.
- [18] D. Kotz and K. Essien, "Analysis of a Campus-Wide Wireless Network," in Proceeding of ACM MobiCom, pp. 107-118, 2002.
- [19] M. Balazinska and P. Castro., "Characterizing Mobility and Network Usage in a Corporate Wireless Local-Area Network," in Proceeding of USENIX MobiSys, 2003.
- [20] IEEE 802.11f, "IEEE Trial-Use Recommended Practice for Multi-Vendor Access Point Interoperability via an Inter-Access Point Protocol across Distribution Systems Supporting IEEE 802.11. Operation," 2003.