Pre-scheduled Handoff for Service-aware and Seamless Internet Access

Kun Xie[1,2], Jiannong Cao[3], Xin Wang[2], Jigang Wen[4]

[1] *College of Computer Science and Electronics Engineering, Hunan University, Changsha, 410082, China*
[2] *Department of Electrical and Computer Engineering, State University of New York at Stony Brook, USA*
[3] *Department of computing, The Hong Kong Polytechnic University, Hong Kong*
[4] *Institute of Computing Technology, Chinese Academy of Science, China*
kun.xie@stonybrook.edu, csjcao@comp.polyu.edu.hk, xwang@ece.sunysb.edu, wenjigang@ict.ac.cn

Abstract

Existing handoff techniques connect a user to a nearby access point whenever it is available along the user's moving path. If the path is not fully Wi-Fi covered, the user may suffer from frequent disconnections. In this work, we propose a pre-scheduled handoff scheme, which determines a quality-aware path so a mobile user can move across an Urban Wi-Fi Mesh network with uninterrupted Internet access at a required data rate. This scheme has several novel techniques. Firstly, by concurrently considering the received signal strength at mobile users and the load of mesh routers (MRs), we design two path finding metrics, *contact weight* and *handoff count*, according to the coverage information. Secondly, to measure the path metric in a practical mesh network with irregular and dynamic MR coverage, we propose a light-weight Bloom-filter-based polling algorithm to discover the mobile stations (called STAs) covered in MR, and a Bloom-filter-based probabilistic algorithm to estimate the degree of coverage overlapping among MRs. Thirdly, based on the path metrics, our proposed scheme supports flexible rate requirements from mobile users through a fully distributed multi-service-class path finding algorithm. Our pre-scheduled scheme considers both mobility and connection requirement, so that the mobile user can continuously access the Internet with the expected data rate. To our best knowledge, this is the first work to ensure seamless Internet access with rate consideration by finding a Wi-

Fi covered geographic path in a practical wireless mesh network. We have carried out extensive simulations, and the simulation results show that the pre-scheduled handoff scheme can provide appealing high data rate to mobile user with minimum handoff delay.

Keywords: Seamless Internet Access, Handoff, Mobility Schedule.

1. INTRODUCTION

Wireless mesh networks (WMNs) have attracted significant attentions in recent years as an avenue to provide cost-effective last-mile network access and wireless Internet services to a large coverage area with a low infrastructure cost [1, 2, 3]. Meanwhile, the bloom of the advanced wireless applications drives the increasing requirements on providing high performance continuous Internet access for mobile users.

In a WMN, a user could access the MRs through a local area network such as 802.11 network, while MRs can communicate with each other through the backbone links such as WiMAX. Similar to wireless local area networks (WLANs), mobile devices in WMNs may experience more frequent handoffs than those in cellular networks because of the shorter transmission range of mesh routers (MR), which makes the performance degradation (e.g.,interrupted connection, packet loss and delay) a more serious problem. To address problems associated with the handoff, many efforts have been made to improve the handoff process. The solutions include shortening the time of L_2 channel scanning [4, 5, 6], optimizing Mobile IP for a better L_3 handoff support [7, 8, 9], improving multi-hop routing to facilitate handoff [10, 11, 12, 13, 14, 15, 16], and other methods from various aspects [17, 18, 19, 20, 21, 22, 23, 24, 25, 26]. These studies generally attempt to increase the Internet access in a best effort and opportunistic manner: providing a best connection whenever it is available along a given path.

However, these opportunistic solutions can hardly provide the mobile users with continuous and high-performance Internet access in WMNs because mobile users may suffer from frequent coverage loss or low quality connection. Based on the experimental studies on vehicular access of existing WiFi networks in urban areas, the authors in [27] observe that the median of the duration of an AP connection is about 13 seconds, while the average inter-arrival time for "associable" APs is about 75 seconds, which implies the long period of disconnections. Several other experimental studies

2

[28, 29, 30] also demonstrate the connectivity intermittence is hardly avoidable in current WiFi networks.

To provide mobile users with a better WiFi-covered path for Internet access, this paper studies a novel problem to guide a mobile user to walk along a good Wi-Fi covered geographic path with well designed handoff strategy. Fig.1 shows a motivation example of our work. Assume a city area is covered by a WMN with several MRs, and a user who is watching an online video showing the stock changes or attending a critical teleconference wants to move from the location MR_1 to MR_4. If the mobile user moves along the shortest geographic path without good Wi-Fi coverage, he would be annoyed by the constant connection interruptions. Alternatively, the mobile user may choose to move along the path $MR_1->MR_2->MR_3->MR_4$ which is fully covered by Wi-Fi and provides him with high quality and continuous video services.
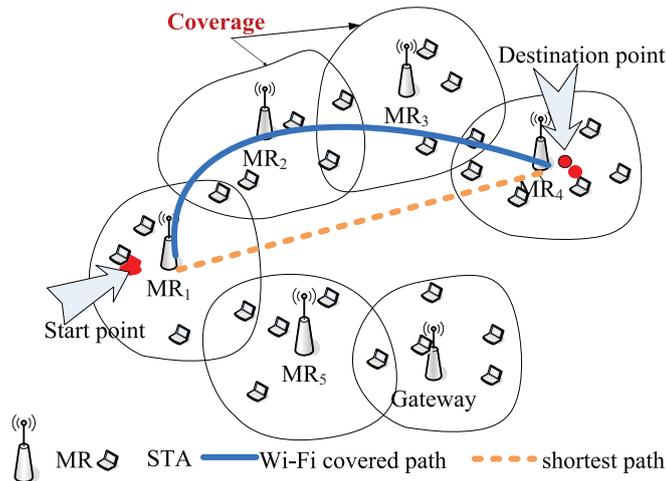
Figure 1: A motivation example.

Obviously, solving the above problem requires the concurrent consideration of geographic path finding and handoff strategy. The geographic path is formed with physical road segments that a human user could walk along. This is totally different from an WMN routing whose goal is to search for a transmission path to deliver packets from the source to the destination without considering the physical road condition, the wireless coverage, and the stability of the connection in the remaining communication time. The major challenge we consider in this paper is to ensure the users to have the

3

Internet access through the MRs on the road over local area networks, not the finding of backbone paths over wide-area links between MRs to transmit the aggregate traffic.

To facilitate the finding of a Wi-Fi covered geographic path, we first define a coverage link by considering both the Wi-Fi coverage and the geographic reachability for the user, and then propose a quality-aware pre-scheduled handoff scheme to determine a path that a mobile user can move across an Urban Wi-Fi Mesh network with uninterrupted Internet access at a required data rate with higher probability. Similar to the path finding service provided by the google map, our scheme provides an option for a user that needs continuous Internet access and certain data rate to find a geographic path that is well covered by the urban WMN. Our pre-scheduled handoff scheme includes the following novel techniques.

First, we propose two novel metrics to facilitate the finding of a road path for a mobile user, *contact weight* and *handoff count*, according to the coverage information about the WMNs. By taking into account both the possible signal strength and the load of an MR in the metric design, the proposed metrics intend to provide an estimate on the quality of data service that a mobile user might experience when moving through the network.

Second, to calculate the contact weight metric in practical WMNs, we first propose a light-weight Bloom-filter-based polling algorithm to discover the STAs covered by an MR, and then propose a Bloom-filter-based probabilistic algorithm to estimate the degree of coverage overlapping among MRs to guide the movement of mobile users. Our proposed contact weight can reflect the strength of received signals and thus the data rate of the network connection. More importantly, our contact weight can be simply estimated without assuming regular domain geometry and can better capture the coverage condition in a complex wireless environment. Our proposed scheme only needs simple Bloom-filter messages to be exchanged among MRs.

Third, to find the path that better meets the quality expectation of a mobile user and to design the strategy of mobility scheduling, we formulate the path finding problem as a link-constrained path optimization problem to select the path that minimizes the handoff frequency while considering the required data rate. The minimization of the number of handoffs also helps a user to find a path with shorter traveling distance under his performance constraints. Moreover, to provide users with flexible service support, we propose a distributed multi-service-class path finding algorithm.

Finally, based on the knowledge of each next MR on the path, we propose

4

a cross-layer connection scheduling strategy, in which both the L_2 and L_3 handoffs are well pre-scheduled which largely reduces the handoff delay.

We have carried out extensive simulations. The simulation results demonstrate that our pre-scheduled handoff scheme can provide appealing high data rate to mobile users with the low handoff delay, which cannot be provided by any other existing work.

It is worth pointing out that the proposed pre-scheduled handoff scheme may be implemented on mobile terminals to supplement the conventional opportunistic handoff scheme [22, 4, 5, 6]. When moving in an area covered by a WMN, a user can choose the handoff mode (pre-scheduled handoff or opportunistic handoff) based on the user's application requirement. A user can choose the pre-scheduled handoff to meet the requirements of critical applications and confirm the path when it is convenient. A mobile terminal equipped with multi-mode handoff schemes is similar to a vehicle equipped with powerful GPS terminal, where a driver could decide if he would follow the GPS guidance based on his need.

The rest of this paper is organized as follows. The related work is presented in Section 2. The system model and problem description is presented in Section 3. The path finding metric is described in Section 4. The path finding algorithm, the mobility scheduling strategy, and the connection scheduling strategy are described in Section 6. We evaluate the performance of the proposed scheme through simulations in Section 7. Finally, we conclude the work in Section 9.

2. RELATED WORK

A complete handoff procedure in WMNs requires the mobility support from both L_2 and L_3 [22]. The L_2 handoff process in IEEE 802.11-based wireless networks can be divided into three steps: scanning, authentication, and reassociation [6]. In L_2 handoff, the delay is mainly resulted from the scanning process which is proportional to the number of channels scanned, and occupies the largest proportion of the entire L_2 handoff delay (more than 90 percent [31]). Extensive studies have been conducted to reduce or eliminate the channel scanning delay during the WMN handoff processes [4, 5, 6].

The L_3 handoff process includes the IP address and routing path updates. As IEEE 802.11 WMNs do not provide sufficient mobility and handoff support, a number of attempts have been made to enable the L_3 handoff in

WMNs. MIPv4 and MIPv6 are the main mobility solutions at the L_3 layer [7, 8]. Although MIP-based handoff schemes have good performance in supporting mobility related traffic delivery in the wired backhaul networks, the performance degrades in multi-hop WMNs. The authors in [9] propose a mobile agent(MA)-based handoff architecture for WMN, assuming each mobile user has an MA residing on his registered mesh router to facilitate the handoff signaling process. Besides above MIP and agent-based L_3 handoff approaches, several L_3 approaches are proposed in WMNs to manage node address change [10, 11] and to modify the multi-hop routing protocol to facilitate the handoff [12, 13, 14, 15, 16].

Our pre-scheduled handoff scheme includes a cross-layer scheduling of L_2 and L_3 connection setup to reduce both handoff delays. In contrast to existing studies, in our scheme, a mobile user is able to make the handoff decision in advance based on the knowledge of each next MR on the path. Therefore, the STA only needs to check one working channel of the next MR and the scanning delay can be largely reduced. Moreover, with early L_2 handoff event triggers, the handoff delay in L_3 can also largely decrease by preparing the L_3 handoff in advance.

There are many other handoff solutions in WMNs. The work in [17] finds that the rayleigh fading and teletraffic parameters play significant roles in evaluating the handoff performance, based on which researchers can study the interaction between physical link layer parameters and higher layer protocol designs. Some recent work in [23, 24] predicts the next AP to connect with and prepares for the handoff in advance to reduce the handoff delay. The work in [18] studies the evolutions of WMN handoff approaches for public safety and disaster recovery networks. Yang et. al in [19] propose a joint handoff and energy management solution for WMNs, in which two distributed handoff schemes are proposed for WMNs with adaptive transmission power and rate. Observing the large handoff delay as a result of the queueing delay and medium access delay at each MR due to the multi-hop transmissions of signaling messages for handoff, the authors in [20, 21] propose channel allocation and channel splitting strategies to reduce the channel access delay. Previous solutions on handoff management in infrastructure-based WMNs mainly assume there exist a single gateway in WMNs. Some recent studies [22, 32, 33] consider inter-gateway handoff and propose new architectures such as IMex [22] to facilitate the inter-gateway handoff management.

In summary, existing handoff solutions only consider the issue of connection scheduling and attempt to maintain network connections in an oppor-

tunistic manner: providing a best connection whenever it is available along a given path. There is little consideration on the throughput improvement and the reduction of network disconnection as a result of terminal movement.

In contrast, this paper studies a novel problem with the goal of facilitating a mobile user to find a Wi-Fi covered geographic path that provides seamless Internet access while at the same time supporting a higher data rate. In addition, this paper concurrently considers the geographic path finding and handoff and proposes a pre-scheduled handoff scheme in which both the mobility and the connection are well scheduled. To the best of our knowledge, this is the first work that proposes a pro-active and pre-scheduled handoff scheme to ensure higher-quality and seamless Internet access.

3. SYSTEM MODEL AND PROBLEM DESCRIPTION

3.1. System Model

We consider a widely accepted architecture for urban WMNs, which has two tiers as illustrated in Fig.2. The access tier connects the client wireless devices (e.g., a wireless laptop in a home, or handheld devices carried by a user on the move) to an MR. The backhaul tier interconnects the MRs to forward traffic to and from the wireline Internet entry points or gateway MRs. Like imesh [34], in this paper, we use OLSR (Optimized Link State Routing) as the routing protocol for backhaul network to route the packets. The term "path" is used in this paper to indicate the path a user moves along, while the term "route" indicates the set of MRs along which packets will be forwarded to reach the destination.
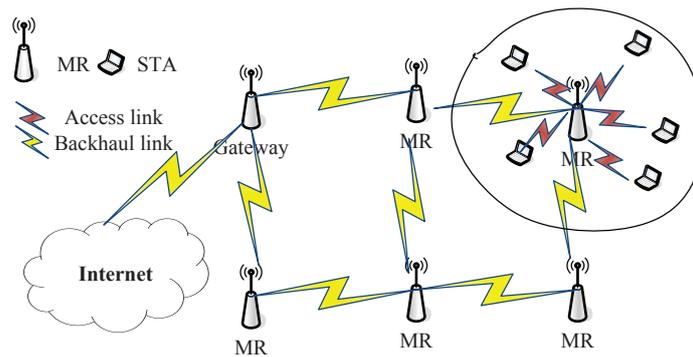
Figure 2: A two-tier WMN.

7

Each MR is equipped with two types of radio, a radio such as 802.11 to serve as an access point for local area network transmissions from mobile users, and a radio such as WiMAX which has a larger capacity and range to relay the traffic in the backbone. Regardless of the actual radio techniques used, access networks in a neighborhood will use orthogonal channels to avoid interference.

Each MR has a limited local coverage area, which is determined by the transmission range and the environment, and can serve only those STAs that reside within its coverage area. A mobile user has to perform handoff to associate with another MR to obtain Internet access when it moves out of the coverage of its current MR and enters the coverage area of another MR.

In this work, any MR in the wireless network can serve as a coordinator to facilitate its associated mobile users to find the optimal moving path. The MR informs a user the path found based on which the mobility of the user and the connection with the network can be pre-scheduled. A user can walk along the path given on the map towards his destination. If a user is willing to change the path, in case of the large rate change, he can be updated with the information on the remaining path by an MR he associates with on the road.

3.2. Problem Description

The problem of this paper: given the starting point and destination point and the required data rate, find out a Wi-Fi covered path with the minimum handoff frequency for the mobile user, so that it can obtain service-aware and seamless Internet access.

The main function of a navigation system is to find a feasible path based on the map of the region of interest. In a general navigation system, different map features are extracted to generate a feasible path and instructions about the surrounding to the user during the actual navigation process. To find a feasible Wi-Fi covered path in this paper, all the map-related data for path finding could be obtained as done in the conventional navigation systems. In order to take advantage of the proposed service to guide users to move along the path, an MR and an STA can determine their positions through localization hardware such as GPS.

In this paper, path, mobility and connection are well scheduled for the mobile users to continuously obtain the required data rate. To facilitate a user to seamlessly access the Internet with a rate expectation, this paper

proposes an efficient fully distributed pre-scheduled handoff scheme by answering the following three questions:

1) How to define a path metric related to the quality of data service a mobile user might experience?

2) How to design the mobility scheduling strategy and how can an MR determine an optimal Wi-Fi covered path for the associated mobile user?

3) What connection scheduling strategies and handoff procedures should be adopted to schedule the mobile user on the path?

In the following sections, we will describe the key techniques in our pre-scheduled handoff scheme.

4. PATH FINDING METRIC

Our study focuses on the handoff scenario, where the major concern is the wireless coverage of the APs or MRs deployed. Our goal is to find a Wi-Fi covered geographic road path for a moving human user to walk towards his destination with well designed handoff strategy. Our problem is thus totally different from a conventional routing problem whose goal is to search for a route to deliver packets along links between routers.

We don't consider backbone transmission links between routers, but instead introduce a *virtual link*, called coverage link, along with two metrics to reflect the quality of network access through an MR at the local area. The coverage link is defined by considering both the wireless coverage and the existence of the road path between two MRs.

Although various routing metrics [35] are proposed in wireless mesh networks to present the quality of a link connecting two routers, such as ETT, ETX [36], WCETT[37], MIC[38], CCM[39], CACU[40], and CMetric [41], they can not be directly applied to our handoff problem. These metrics do not consider the physical road conditions. Also, the quality of the virtual link cannot be easily measured.

4.1. Coverage Link

Definition 1. A *coverage link* between two MRs (MR_a and MR_b) is defined as the shortest feasible link that connects MR_a and MR_b and satisfies the following two conditions.

- **Condition 1** From the coverage perspectives, MR_a and MR_b are neighbors. An MR is called a neighbor of the reference MR if its coverage overlaps with that of the latter's.

9

- **Condition 2** From the geographic perspectives, a user can go from MR_a to MR_b physically without obstacles on the road. According to the map-related information, we can easily identify that whether there exists a geographic route or not.

Therefore, a coverage link is defined by considering both the Wi-Fi coverage and the physical reach-ability for a moving user. A user moving along a coverage link has a high probability of accessing the Internet without communication interruption. In the remaining of the paper, we simply call the coverage link as a link, and we mainly care about the performance along the coverage link. As far as we know, this paper is the first that introduces the notions of coverage neighbor and coverage link and takes them into account in the handoff schedule design.

Based on the coverage information, in this section, we introduce two performance metrics (Contact Weight and Handoff Count) to effectively capture the quality of data service that a mobile user might experience when moving through the network.

4.2. Contact Weight

The data rate a mobile user can obtain is based on the strength of the received signal. Generally, according to the wireless channel fading model, the signal strength decreases if the transmission distance increases. A mobile user that is far away from its associated MR and on the boundary of the coverage area will have a lower rate.

(a) light degree (b) heavy degree
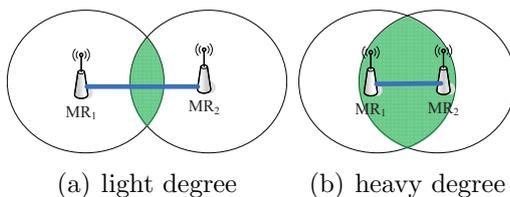
Figure 3: The degree of coverage overlapping.

As shown in Fig.3, the degree of coverage overlapping of two neighboring MRs (MR_a and MR_b) can reflect the strength of the received signal on the link $(a - b)$. If two adjacent MRs has a higher degree of coverage overlapping, it implies either the distance between these two MRs is smaller or the transmission channel is good, and a mobile user moving along the link between these two MRs can associate with an MR with a better signal during

10

the handoff to receive a higher average data rate with a higher probability. Although the channel condition of a link can vary over time, before a mobile is associated with a MR, there is no physical wireless link to measure. Therefore, we only estimate the signal strength based on the estimation of the coverage overlapping between two MRs.

Besides the received signal strength, the load of MRs in WMNs may also contribute to the data rate a user experiences. Finding a path only based on the signal strength without considering the load condition of MRs is not sufficient, as otherwise the MRs on some popular paths may get overloaded which would result in the rate reduction of each user.

Therefore, taking into account both the signal strength received at mobile users and the load factor in WMNs, we propose a novel metric *contact weight* which can be calculated as

$$ContactWeight_{a-b} = \frac{Overlapping_{a-b}}{1 + \max\{\hat{L}_{a_{avg}}, \hat{L}_{b_{avg}}\}} \tag{1}$$

where $Overlapping_{a-b}$ denotes the degree of coverage overlapping of two neighboring MRs (MR_a and MR_b), $\hat{L}_{a_{avg}}$ and $\hat{L}_{b_{avg}}$ denote the average load of MR_a and MR_b, respectively. In Section 5.1, we will introduce our method to evaluate the degree of coverage overlapping of two MRs.

Obviously, a larger degree of coverage overlapping with a smaller load may make the STA experience higher data rate and result in the value of the ContactWeight to be larger. Therefore, a path with a bigger ContactWeight may provide the mobile user with a better transmission rate.

Before the handoff execution, the user associates with MR_a, and the load of MR_a impacts the experience of the user. Similarly, the load of MR_b impacts the user's experience after the handoff execution. In our model, the Wi-Fi path for a mobile user should consider the minimum data rate the user requires. As the minimum data rate achieved along a link depends on the highest load of MR_a and MR_b, to reflect the user's experience of Internet access along a link, this highest load is defined as the link load in Eq.(1). In (1), note we add 1 in the denominator to avoid its being zero when no load is generated at either MR_a or MR_b.

4.3. Handoff Count

During handoff, a connection of the mobile user may experience a sudden decrease of throughput or even outage due to an unavoidable high handoff

delay as a result of channel scanning, authentication and re-association procedures. In addition, a handoff process may involve other message overhead to maintain the connection between an MR and its STAs. To minimize the performance degradation and message overhead, an optimal path selection scheme should minimize the handoff frequency on the path. Extending the concept in [17], we define the *handoff count* as the number of handoff operations along the moving path of a mobile user. As a handoff happens when a user changes its association from one MR to another MR with the two forming a coverage link, the *handoff count* of a path is equal to the number of links on the path and also the hop count of the backbone meshed network.

To better meet the user's rate requirement, our algorithm will look for a path with good Wi-Fi coverage and light load. We will present our algorithm based on the proposed metrics, contact weight and handoff count, in Section 6. Our algorithm is also designed to achieve low handoff frequency.

5. Metric calculation

According to (1), given two adjacent MRs (MR_a and MR_b), to calculate the contact weight of the link ($a - b$), we should first evaluate the degree of coverage overlapping of these two MRs, and then the average load of MR_a and MR_b. In this section, we present our method of calculating these required parameters.

5.1. Calculating the degree of coverage overlapping

In an urban WMNs, terrain, foliage, weather, obstacles, node movement, and the presence of emitters of electromagnetic radiation all affect the transmission range of MRs, so the coverage of an MR follows an irregular geometric shape. We can not design a geometry-related solution to discover coverage information by assuming that the wireless coverage area is contiguous and with a regular shape such as circular, sector, etc. Therefore, accurately discovering an MR's coverage and evaluating how much the coverage is overlapped with that of a close-by MR are challenging. In this paper, we use a trick to estimate the coverage of an MR. More precisely, we use the set of STAs covered by an MR to denote the MR's coverage.

Let S_a and S_b be the STA sets covered by MR_a and MR_b. The size of S_a and S_b are denoted as n_a and n_b respectively. The size of $S_a \cup S_b$ is denoted as $n_{a \cup b}$. Therefore, the degree of coverage overlapping of this pair of MRs

can be calculated and normalized as follows.

$$Overlapping_{a-b} = \begin{cases} \frac{(n_a + n_b - n_{a \cup b})}{(n_a + n_b)} & n_a + n_b \neq 0 \\ 0 & otherwise \end{cases} \qquad (2)$$

where $Overlapping_{a-b}$ is bounded between 0 and 1. Such a normalization can provide a better estimate of the degree of coverage overlapping, and avoid the bias when the number of STAs associated with an MR is too low or too high.

To calculate the degree of coverage overlapping of two adjacent MRs in practical WMNs, we need to address the following key issues:

1) How does an MR discover its coverage and transmit the coverage information to the MRs close-by?

2) How does the MR quantitatively estimate the degree of coverage overlapping among MRs?

In the following subsection, we propose two algorithms to address these issues.

5.1.1. Light-weight Bloom Filter Based Polling Algorithm
  • Coverage Representation
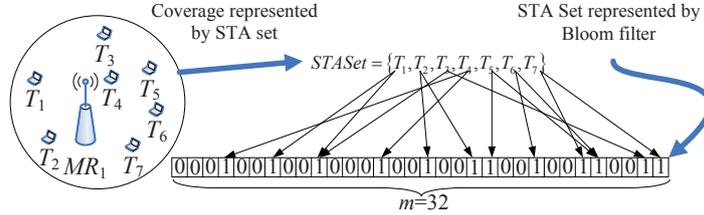
Figure 4: MR coverage representation.

In wireless networks, the bandwidth is limited, especially when the network has a large number of users competing in channel access. To reduce the signaling overhead, the coverage representation should be efficient.

A Bloom filter[42], which is known to be a space-efficient data structure, has a great potential for use in a distributed system to share information about available data. A Bloom filter uses an m-length string to represent a data set with a size of n. To represent an element by a Bloom filter, the element is hashed by k functions, and the corresponding bit from each hash

is set to 1 in the Bloom filter [42]. The source element cannot be discovered from the Bloom filter and the privacy of the element is protected. Therefore, we apply Bloom filter to represent the coverage information and exchange the simple Bloom filter strings among neighboring MRs. In the example shown in Fig.4, the size of the STA set is $n = 7$, the number of hash functions is $k = 2$, and the length of the Bloom filter string is $m = 32$.

- Coverage Discovery

Although an MR is aware of its associated set of STAs, it is hard for the MR to find out the STAs that are also within its coverage area but are currently associated with other MRs and thus turn their radios to channels taken by their individual associated MRs. Fig.5 illustrates the challenge of the coverage discovery problem. As an example, MR_A needs to discover the STA set comprising of $T_1, T_2, T_3, T_4, T_5, T_6, T_7$. Although T_6 and T_7 are in the coverage of MR_A, they are associated with MR_B and transmit packets through Ch_2. Thus these two STAs can hardly be discovered by MR_A.
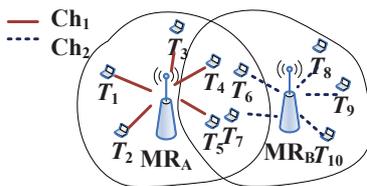
Figure 5: Coverage discovery problem.

Under the standard Wi-Fi protocol, an STA can scan all channels to find the best MR to associate with during its handoff. Requesting all STAs to scan different channels actively and send messages to each MR within the coverage would consume a lot of bandwidth.

Alternatively, to get the statistics on the coverage condition, an MR can periodically poll the channels of the close-by MRs, and an STA receiving the polling with a good signal strength can return its ID. An MR could suspend its communications with the associated STAs using a NAV window. To avoid the communication suspension, the MR can use another radio to send the polling message and wait for the responses of the STAs. Although STAs may respond directly to the query of an MR, this again, could potentially lead to higher overhead and consume lots of bandwidth.

Instead, we propose a light-weight Bloom-filter-based polling for MR to discover the STAs associated with other MRs, as shown in Algorithm 1. To solve the coverage discovery problem as shown in Fig.5, when MR_A wants to collect the information on STAs which are covered by MR_A but associated with MR_B, it will send out a query message on the working channel Ch_2 of the neighbor MR_B to learn the number of STAs, as shown in Fig.6(a).

The channels of neighboring MRs are generally orthogonal to avoid interference. If the channel can change over time, the change at an MR will be announced to its neighbors. The polling can be done when there is a need to look for the path and there is a time duration from the last update of the STA knowledge. It will not be performed frequently. This needs a small protocol support at the MAC layer along with the association part for users to respond. To avoid interfering with existing transmissions, a request can be first sent to the neighboring MR, which will reserve a time duration using NAV window for the querying MR.

Along the polling message, the MR_A also sends out a schedule with m slots for responses where m is the length of the Bloom filter, and initializes the Bloom filter SBF_{Ch_2} to zero. To apply the Bloom filter to represent an STA set in the coverage of an MR, we let STAs to fill the Bloom filter intelligently through the analog tone transmission in successive m time slots. The STA transmissions in a sequence of slots can be similar done by a TDMA kind of procedure in general. An STA receiving the query selects the time slots for responding based on the hash results of its ID, as shown in Fig.6(b) in the paper. The STA will broadcast a simple tone signal during the time slot l if one of the hash values $h_p[ID](1 \leq p \leq k)$ is l. MR_A sets the bit of SBF_{Ch_2} to one if it receives the tone signal in a corresponding slot. The slot level synchronization has been considered in current 802.11. The query messages will further facilitate different STAs to coordinate in sending responses in different slots. The implementation of polling and response is simple. Compared to complete message feedbacks from each covered STA, the simple beacons have much lower signaling overhead.

A Bloom filter is a bit string often utilized to represent a data set. The length of bloom filter, m, is usually determined by the number of elements in the data set to control the probability of bit setting as a result of the hashing operation. From literature studies, a Bloom filter can best represent a data set when the probability of bit setting is 0.5. To apply the Bloom filter to represent an STA set in the coverage of an MR, we can first estimate the size of STA set based on the past traffic and then determine the m accordingly.

(a) Message Sequence Diagram   (b) Time slot calculation
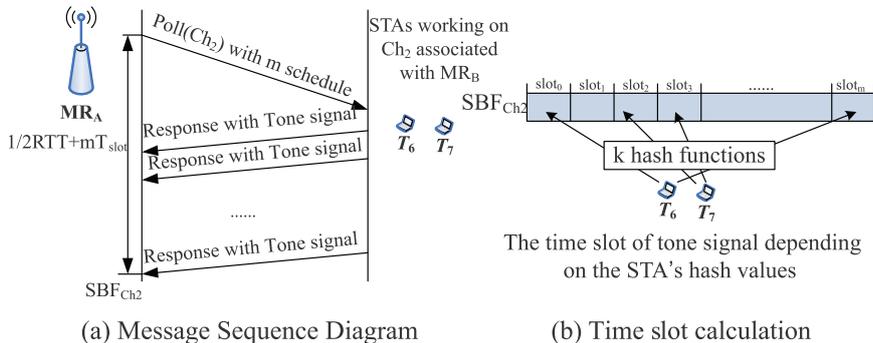
Figure 6: Bloom filter based polling for STA discovery.

The complete procedure for an MR_A to discover its covered STA set can be described in Algorithm 1. The MR_A first takes a Bloom filter string SBF_A to represent the STAs associated with it. It then polls the other MRs' working channels one by one exploiting use of the light-weight Bloom filter. During the query process, the STAs located in its coverage area but associated with other MRs can be discovered and the STAs working on the channel $ch(j)$ are represented by $SBF_{ch(j)}$.

From [42], we know that the union of two sets can be directly represented by employing the OR operation on the Bloom filter strings that respectively represent the two sets. Therefore, MR_A can merge SBF_A with all the other $SBF_{ch(j)}$ through the "OR" operation to form the SBF_A to represent all STAs in its coverage area.

From Fig.6, the MR will spend a time period of $1/2RTT + mT_{slot}$ to discover the number of STAs associated with a neighboring MR. To obtain the coverage information in a dynamic wireless environment, MR can execute the light-weight Bloom-filter-based polling algorithm periodically. Because the algorithm is designed based on Bloom filter, the time and message overheads of our coverage discovery procedure are very low, which makes it light weight and easy to implement over the off-the-shelf WiFi products, following the signal diagram in Fig.6.

To calculate the $ContactWeight_{a-b}$ of $link_{a-b}$ based on Eq(2), after an MR obtains SBF_a and SBF_b corresponding to MR_a and MR_b, it needs to find the sizes of the STA set of MR_a, MR_b, and their union set, which are denoted as n_a, n_b, and $n_{a \cup b}$ respectively. In the next subsection, we propose a probabilistic algorithm to estimate the size of STA set based on the bit strings of Bloom filter exchanged among MRs.

16

Algorithm 1 Bloom filter based polling algorithm

Input: MR_A and its neighboring MRs.
Output: SBF_A which represents all the STAs covered by MR_A.
    Part 1: Represent the STAs associated with MR_A
1: **for** each STA_i associated with MR_A **do**
2:     The k hash functions are applied to $STA_i's$ $ID(i)$ to compute the corresponding k locations of the SBF_A, and the corresponding k bits are all set to 1.That is, for $1 \leq y \leq k$, let $SBF_A[h_y(ID(i))] = 1$.
3: **end for**
    Part 2:Represent the STAs associated with other neighboring MRs
4: **for** each MR_j in the neighborhood of MR_A, and its working channel is $ch(j)$ **do**
5:     MR_A sends out a $poll(ch(j))$ message and a schedule with m time slots for response
6:     An STA receiving the poll locally hashes its ID by using the k hash functions, and broadcasts a simple signal tone during the time slot l hashed by its ID.
7:     MR_A sets the bit of $SBF_{ch(j)}[l]$ to 1 if it receives the tone signal in time slot l.
8: **end for**
    Part 3: Merge Bloom filters
9: $SBF_A = SBF_A \cup \bigcup\limits_{ch(j) \ of \ all \ neighbor \ MR_j} SBF_{ch(j)}$

5.1.2. Bloom Filter Based Probabilistic Algorithm

From Algorithm 1, SBF represents STA sets by utilizing k hash functions to set the k bits of SBF to 1. Even though some STAs may share the same bits, the number of bits set to 1 in the Bloom filter with the increase of the number of STAs. Therefore, we propose a probabilistic algorithm to estimate the size of STA set as shown in Theorem 1 by utilizing the occupancy information of the Bloom filter.

Theorem 1 Suppose SBF_{S_1} represent the STA set S_1 of MR_1 using a Bloom filter of length m and k hash functions, and the number of bits set to 1 in the SBF_{S_1} is s. Then the number of STAs in S_1 is $n = \log\left(1 - \frac{s}{m}\right) / \left(k \times \log\left(1 - \frac{1}{m}\right)\right)$.

Proof. We assume that there are totally n STAs in S_1. After being hashed into the SBF_{S_1}, the probability that a specific bit is still 0 is $p(n) = \left(1 - \frac{1}{m}\right)^{kn} \approx e^{-kn/m}$. The probability that a specific bit is set to 1 can be expressed as $1 - p$. After representing n STAs, the total expected number of bits set to 1 is $\hat{S}(n) = (1-p) \times m = \left(1 - (1-1/m)^{kn}\right) \times m$. And we

further have

$$\hat{S}(n) = (1 - p) \times m = \left(1 - (1 - 1/m)^{kn}\right) \times m$$
$$\Rightarrow 1 - (1 - 1/m)^{kn} = \hat{S}(n)/m \tag{3}$$
$$\Rightarrow n = \log\left(1 - \hat{S}(n)/m\right)/\left(k \log\left((1 - 1/m)\right)\right)$$

Let $s = \hat{S}(n)$, the proof completes. ∎

Thus, to calculate the degree of coverage overlapping of a pair of MRs as shown in Eq(2), we first employ the OR operation on the Bloom filters to estimate the cardinality of the set union. Then, we can use the estimated set sizes of $|S_1|$, $|S_2|$, and $|S_1 \cup S_2|$ to obtain the degree of coverage overlapping of $\frac{|S_1| + |S_2| - |S_1 \cup S_2|}{|S_1| + |S_2|}$, where $|S_1| + |S_2| - |S_1 \cup S_2| = |S_1 \cap S_2|$.

The cost on calculating the degree of coverage overlapping is low. Given a Bloom filter, to estimate the size of STA set represented by the Bloom filter, the cost for counting the number of "1" in a Bloom filter of length m is $O(m)$. Also, our proposed probabilistic estimation is performed based on simple binary operations of Bloom filter strings, which can be easily implemented on well-designed hardware at the low complexity.

Following we will analyze the accuracy of our proposed probabilistic estimation algorithm. The error margins can be calculated for given intervals of set sizes following Theorem 3.

Theorem 2 Suppose SBF_{S_1} represent the STA set of MR_1, which is denoted as S_1 and includes n STAs. SBF_{S_1} is a Bloom-filter of length m, which is formed with k hash functions. After representing n STAs, the expected number of bits that are set to 1 in SBF_{S_1} is denoted by $\hat{S}(n)$. Then the probability for the number of "1" bits to be more than $(1 + \delta)\hat{S}(n)$ is $p\left(TotalBit > (1 + \delta)\hat{S}(n)\right) \leq e^{-\hat{S}(n)\delta^2/3}, (0 < \delta < 1)$. The probability of number of bits which are set to 1 to be less than $(1 - \delta)\hat{S}(n)$ is $p\left(TotalBit < (1 - \delta)\hat{S}(n)\right) \leq e^{-\hat{S}(n)\delta^2/2}, (0 < \delta < 1)$.

Proof. We assume that n STAs are hashed into the SBF. The probability that t bits are set to 1 is a binomial random variable, and can be expressed as $p(TotalBit = t|n) = C_m^t\left((1 - p(n))^t p(n)^{m-t}\right)$. Using Chernoff bound [43], we can obtain the upper bound and lower bound of the probabilities, that is, $p\left(TotalBit > (1 + \delta)\hat{S}(n)\right) \leq e^{-\hat{S}(n)\delta^2/3}, 0 < \delta < 1$ and $p\left(TotalBit < (1 - \delta)\hat{S}(n)\right) \leq e^{-\hat{S}(n)\delta^2/2}, 0 < \delta < 1$. The proof completes. ∎

Theorem 3 Suppose a Bloom filter SBF_{S_1} represent STA set S_1 of MR_1 with the filter length m and k hash functions, and the number of bits set to 1 is s. For any n_l and n_u, the number of STAs hashed into SBF_{S_1} is denoted by n and lies in the range (n_l, n_u) with a probability of at least $1 - e^{-\hat{S}(n_u)\delta_u^2/2} - e^{-\hat{S}(n_l)\delta_l^2/3}$, where $(1 + \delta_l)\,\hat{S}(n_l) < s$ and $(1 - \delta_u)\,\hat{S}(n_u) > s$.

Proof. We prove the theorem in the following two steps.

If the number of STAs $n \le n_l$, then we have $p\,(TotalBit \ge s|n) \le p\,(TotalBit \ge s|n_l) = \sum\limits_{t=s}^{m} \left(C_m^t \left((1 - p\,(n))^t p(n)^{m-t} \right) \right)$, which shows that the probability of $p\,(TotalBit \ge s|n)$ decreases when s increases. We choose n_l and δ_l which make $(1 + \delta_l)\,\hat{S}(n_l) < s$. Following Theorem 2, we obtain $p\,(TotalBit \ge s|n_l) < p\left(TotalBit > (1 + \delta_l)\,\hat{S}(n_l)\,|n_l \right) \le e^{-\hat{S}(n_l)\delta_l^2/3}$. Therefore, we can obtain $p\,(TotalBit \ge s|n) < e^{-\hat{S}(n_l)\delta_l^2/3}$.

If $n \ge n_u$, we can obtain $p\,(TotalBit \le s|n) \le p\,(TotalBit \le s|n_u) = 1 - \sum\limits_{t=s}^{m} \left(C_m^t \left((1 - p\,(n))^t p(n)^{m-t} \right) \right)$, which shows $p\,(TotalBit < s\,|n)$ increases as s increases. Similarly to the above proof for $(n \le n_l)$, we choose n_u and δ_u to make $(1 - \delta_u)\,\hat{S}(n_u) > s$, and we can obtain that $p\,(TotalBit \le s|n_u) < p\left(TotalBit < (1 - \delta_u)\,\hat{S}(n_u)\,|n_u \right) \le e^{-\hat{S}(n_u)\delta_u^2/2}$.

Therefore, we obtain $p\,(TotalBit \le s|n) < e^{-\hat{S}(n_u)\delta_u^2/2}$.

Combining $(n \le n_l)$ and $(n \ge n_u)$, we can conclude that the number of STAs (denoted by n) hashed into the SBF_{S_1} lies in the range (n_l, n_u) with a probability of at least $1 - e^{-\hat{S}(n_u)\delta_u^2/2} - e^{-\hat{S}(n_l)\delta_l^2/3}$. The proof completes. ∎

5.2. Calculating the average load of an MR

In a WMN, the load of an MR is not only impacted by the users it serves, but also the load incurred for it to forward the traffic for other MRs in the backhaul network. However, as introduced in the system model in Section 3, the traffic from MRs is relayed with the radio of large capacity and range such as WiMAX in the backhaul network. In this work, we focus on the handoff in local area networks, and assume that the backhaul network has enough bandwidth. Therefore, to determine the load impact on the MNs' rate, we only consider the traffic transmitted using the local-area network resources.

We define the load L_a of MR_a to be the total load generated by all users associated with MR_a, expressed as $L_a = \sum_{u \in U_a} (t_u)$, where U_a denotes the

STA set associated with MR_a and t_u represents the rate of the total data transmitted to and from the mobile user u. To obtain the load L_a, an MR can track the total traffic generated by its associated users.

User data traffic is often bursty which makes the load on the MRs dynamic. To better estimate the load of an MR, we take EWMA (exponentially weighted moving average) to calculate the average load:

$$\hat{L}_{a_{avg}} = (1 - \alpha) \, \hat{L}_{a_{avg}} + \alpha L_a \tag{4}$$

where α is the weight in EWMA.

The average load of an MR will be applied to facilitate the path finding of a user and balance the network load. Each MR periodically calculates and updates the average load by using (4). If the load change of an MR exceeds certain level, the MR will broadcast the updated load with a small TTL through the backhaul mesh link to its neighboring nodes. If the load of an MR changes dramatically and cannot serve a mobile user at the user's expected data rate, the mobile user can inform its current associated MR to initiate an new path search to avoid large performance degradation.

6. PATH FINDING ALGORITHM AND MOBILITY SCHEDULING

6.1. Path Finding Problem Formulation

The aim of our work is to facilitate a mobile user to find a moving path that allows for seamless Internet access while supporting higher transmission quality. A user may inform his closest MR the required data rate according to the desired application quality. MR will return a path that satisfies the rate requirement. For example, a streaming application generally has some rate expectation. On the other hand, streaming sources normally support graceful rate adaptation which is very useful when the allowable transmission bandwidth changes.

Instead of meeting the rate requirement of each user, for the simplicity of system implementation, we consider supporting CoS (class of services) mapping between the available data rate and the quality-level for transmission. Specifically, if an application with the desired quality level A_i requires the data rate to be larger than O_i, this requirement can be mapped to a service class C_i, as expressed in (5).

$$A_i \to C_i \to O_i, \tag{5}$$

where O_i is the minimum data rate of the service class C_i. Based on the CoS mapping, an MR can construct a path table with different paths corresponding to different CoS classes. A mobile user will select the path which satisfies the rate requirement according to application's quality expectation of this user.

Based on the proposed path finding metrics including the contact weight and the handoff count of path, the path finding problem can be formulated as a link-constrained path optimization problem, expressed as

$$
\begin{aligned}
\text{Minimize} \quad & \text{handoff count } (path(s,d)^{C_i}, path(s,d)^{C_i} \in P(s,d)) \\
\text{s.t.} \quad & ContactWeight(link(x,z)^{C_i}) > w_{th_i} \forall x, z \\
& link(x,z)^{C_i} \in path(s,d)^{C_i}
\end{aligned}
\tag{6}
$$

where s and d denote the beginning MR and end MR that cover the starting point and destination point of the mobile user, $P(s,d)$ is the set of available Wi-Fi covered paths. $path(s,d)^{C_i}$ is a path from s to d belonging to the class C_i. w_{th_i} is the minimum contact weight requirement of class C_i. w_{th_i} can be determined based on O_i corresponding to the class C_i. In practice, we can use statistics results from experiments as a reference to determine the required contact weight for a given data rate requirement. In Section 7, we will show a way to utilize cumulative distribution function (CDF) as the statistic tool to obtain the relationship of the value of contact weight and data rate, based on which, the w_{th} is set according to the user's requirement.

The handoff count on a path depends on the number of links on the path. Thus, the path finding problem is to find a path with the minimum number of links with each link satisfying the user's rate requirement. Therefore, the path found may not be long and the mobile user may prefer to move on.

6.2. Problem Transformation

The link constraint $ContactWeight(link(x,z)^{C_i}) > w_{th_i}$ considerably complicates the problem in (6), which makes it difficult to directly solve the problem. To simplify the problem and look for the solution, we update the link cost of $link(x,z)^{C_i}$ according to different class's rate requirement as follows:

$$
Linkcost\left(link(x,z)^{C_i}\right) = \begin{cases} 1 \ ContactWeight\left(link(x,z)^{C_i}\right) \geq w_{th_i} \\ +\infty \ \text{otherwise} \end{cases}
\tag{7}
$$

where the cost $Linkcost\left(link(x,z)^{C_i}\right)$ between two MRs x and z is set to 1 if the link's contact weight is not less than the contact weight threshold w_{th_i};

otherwise, it is set to $+\infty$. By updating the link cost in (7), the problem in (6) can be translated to the one of finding a path with the minimum accumulated cost of all links as

$$Minimize \ path \cos t(path(s,d)^{C_i}, path(s,d)^{C_i} \in P(s,d))$$
$$s.t. \ path \cos t(s,d)^{C_i} = \sum_{link(x,z)^{C_i} \in path(s,d)^{C_i}} Linkcost\left(link(x,z)^{C_i}\right) \quad (8)$$

6.3. Multi-Service-Class Path Finding Algorithm

To find a well Wi-Fi covered path to provide user with seamless and high performance communications and Internet access, the path finding algorithm should meet the following requirements: 1) To flexibly support requirements from different mobile users, the path finding algorithm should be able to find multiple well covered Wi-Fi paths for multiple service classes. 2) In a wireless network, sometimes it may not be possible to guarantee a rate absolutely. A user may choose to accept the highest available rate when his desired rate cannot be ensured. The path finding algorithm should find an alternate path for the user to seamlessly access the Internet at a lower quality level.

According to the requirements, we propose a Multi-Service-Class Path Finding algorithm (MSCPF) to support diverse quality requirements from mobile users, as shown in Algorithm 2. MSCPF can be distributively executed and includes two parts as follows.

In part one, an MR constructs its path table for different service classes. Firstly, each MR calculates the cost of its outgoing coverage links by using (1). Secondly, MR updates the link cost according to different class's rate requirement in (7). Finally, the distributed Bellman-Ford shortest path algorithm [44] is implemented at each MR using these newly updated costs. Each MR_k $(k \in 1,...,N)$ updates its cost towards the destination as

$$P_k^{C_i} = \min_{j \in N(k)} \left(link \cos t(link(k,j)^{C_i}) + P_j^{C_i}\right) \quad (9)$$

where $P_k^{C_i}$ denotes the cost from node k to the destination, $N(k)$ denotes the set of neighboring nodes of MR_k, and $link \cos t(link(k,j)^{C_i})$ denotes the updated cost between node k and node j calculated in Step 3.

In part two, the mobile user informs its closest MR the required quality level, and the MR will determine the class C_i accordingly. If MSCPF algorithm cannot find a path that meets the user's required quality, to make the path finding process more robust, MSCPF can provide the user an alternate path corresponding to a lower class of service at a lower expected rate.

The worst-case computational complexity of calculating the costs at each node in path calculation step is $O(N^2)$ since it requires two nested loops, and each has a maximum length of $N - 1$ to calculate all the nodes.

Algorithm 2 Multi-Service-Class Path Finding algorithm (MSCPF)

Part one: QoS Path construction

1. **Link cost calculation**:Each MR_x $(x \in 1, ..., N)$ calculates the degree of coverage overlapping o_{xz} of its outgoing link (x, z) by using (2) and the average load on the MR_x (denoted as $\hat{L}_x$)by using (4) where $z \in N(x)$ and $N(x)$ denotes the set of neighboring nodes of MR_x. After MR_x obtains $\hat{L}_x$, it exchanges the load information with its neighboring nodes. The cost of the (x, z)-th link is calculated as $ContactWeight(link(x, z)) = \frac{o_{xz}}{1 + \max\{\hat{L}_x, \hat{L}_z\}}$ by using (1).

2. **Update the link cost**: Let $link\,cost(link(x, z)^{C_i})$ denote the updated cost of link (x, z) for class C_i, and w_{th_i} be the contact weight threshold of this class. After we have $ContactWeight(link(x, z)^{C_i})$, the updated cost $link\,cost(link(x, z)^{C_i})$ can be obtained according to (7).

3. **CoS Path calculation**: Apply distributed Bellman-Ford algorithm to build the path table with different paths corresponding to different classes based on the updated link costs. Each MR_k $(k \in 1, ..., N)$ executes the iteration $P_k^{C_i} = min_{j \in N(k)}(link\,cost(link(k, j)^{C_i}) + P_j^{C_i})$, where $N(k)$ denotes the set of neighboring nodes of MR_k, $P_j^{C_i}$ represents the latest estimate of the minimum accumulated cost of the C_i-class-path from MR_j to the destination, and $link\,cost(link(k, j)^{C_i})$ is the updated cost from MR_k to MR_j calculated in Step 2.

Part two: Path selection

1. **CoS class identification:** The closest MR determines the class C_i according to the quality requirement of the user.

2. **Path finding:** MR looks for the path table and checks if there exist one path for this class C_i. If yes, return this path; otherwise go to Step 3.

3. **Update the CoS class:** If C_i is the lowest level class supported, return "No path"; otherwise $C_i = C_{i-1}$ and go to Step 2.

6.4. Mobility and Connection Scheduling Strategies

When a user needs to move from a starting point to a destination point, the user first informs the MR closest about his quality requirement. After running the MSCPF algorithm, the MR informs the user the path found based on which the mobility of the user and the connection with the network can be pre-scheduled. After selecting the Wi-Fi covered path for a mobile

user, the moving direction of the user is scheduled according to the direction of the sequence of links on the path.

To achieve a small handoff delay and packet loss, we concurrently consider L_2 and L_3 handoffs with the support of a well-scheduled handoff strategy. In our design, with the knowledge of the links thus each next MR in advance, the connection schedule for each link can be implemented in a simple way. When moving from MR_A to MR_B, an STA only needs to check the signal strength of the channel of MR_B when the signal received from the MR_A is not good, and the handoff is triggered and the association is changed from MR_A to MR_B when the signal strength from MR_B is higher than a threshold. In our mechanism, an STA knows its next MR, MR_B, in advance. It thus only needs to check the working channel of MR_B, which helps to significantly reduce the scanning delay of L_2. We call such a connection strategy in this paper a *scheduled channel check*.

With our pre-scheduled handoff, the L_2 scanning delay is significantly reduced and a handoff is triggered at L_2 when the signal strength experiences a significant change between two MRs. We also use the signal strength change to trigger the L_3 handoff. When the signal strength indicator, which also functions as the L_3 handoff trigger, exceeds a certain threshold, the L_2 handoff will be triggered, and consequently an L_3 handoff will be triggered. The cross-layer design by using L_2 information benefits the overall handoff performance.

7. SIMULATION

7.1. Simulation Setup

We implement our pre-scheduled handoff scheme by using the simulator NCTUns [45]. Our channel follows the Rayleigh fading model. The maximum transmission range of the access radio and the backhaul radio of each MR are set to 250 meters and 500 meters respectively. The access radios of the close-by MRs are tuned to orthogonal channels. All the MRs connect to a gateway through wireless backhaul links, and the gateway connects to an external server through a wired connection. As described in the system model in Section 3, in urban WMNs, we can use the radio with large capacity and range such as WiMAX to form the backhaul network. We assume that the backhaul network between an MR and the gateway and the wired link between the gateway and the external server can always provide enough

24

bandwidth for all the traffic, whereas the wireless access network is the bottleneck in terms of the data rate. A mobile node (MN) receives packets only from the MR it is associated with. The received data rate is used as a performance metric to evaluate different handoff schemes and measured at the MN. In all simulations, we use UDP traffic instead of the TCP traffic to decouple the transmission behavior from the effect of congestion control. To evaluate the impact of various factors on the rate of an MN, we let the server transmit data at the maximum transmission rate allowed by the current wireless medium. To achieve this, we set the traffic mode of the server to be "ttcp" in NCTUns. Under this mode, the external server can transmit UDP packets to the MN as fast as it can. For all the other STAs associated with MRs, however, we configure their traffic mode to be "stg" instead of "ttcp". The stg mode would ensure the external server to transmit UDP packets to an STA at a constant transmission rate, which would allow us to better control the background traffic load. To vary the total traffic load of an MR, we change the packet length thus the data rate from the server to each STA. The simulation includes two parts.

Part 1: Data collecting and training. In this part, we examine how the degree of coverage overlapping, the load of MRs, and handoff strategy impact the received rate of MN in a linear network. We vary network conditions to get different contact weights, and obtain simulation results of different data rates under various contact weights. Then, according to the probability theory and statistics, we build the benchmark which describes the relationship of data rate and the value of contact weight.

Part 2: Guiding user movement with the training data. We perform simulations with MRs randomly deployed in an area of 4000 m × 4000 m. We use the benchmark built in the first part as the base to guide the mobile user movement in the network to obtain its desired rate.

7.2. Simulation Results (Part 1)

In this part, we carry out the simulations in a linear network to obtain a large number of simulation results to train the benchmark. In the linear network, some MRs are deployed along a line of 2000 m with the same distance between MRs, and these MRs connect with the gateway over one or a few hops. In the linear network, the STAs are uniformly distributed in the network. By changing the location of MRs and the STA density, we can change the value of the degree of coverage overlapping of each pair of neighboring

MRs and the total load generated by STAs, and thus change the value of weight contact.
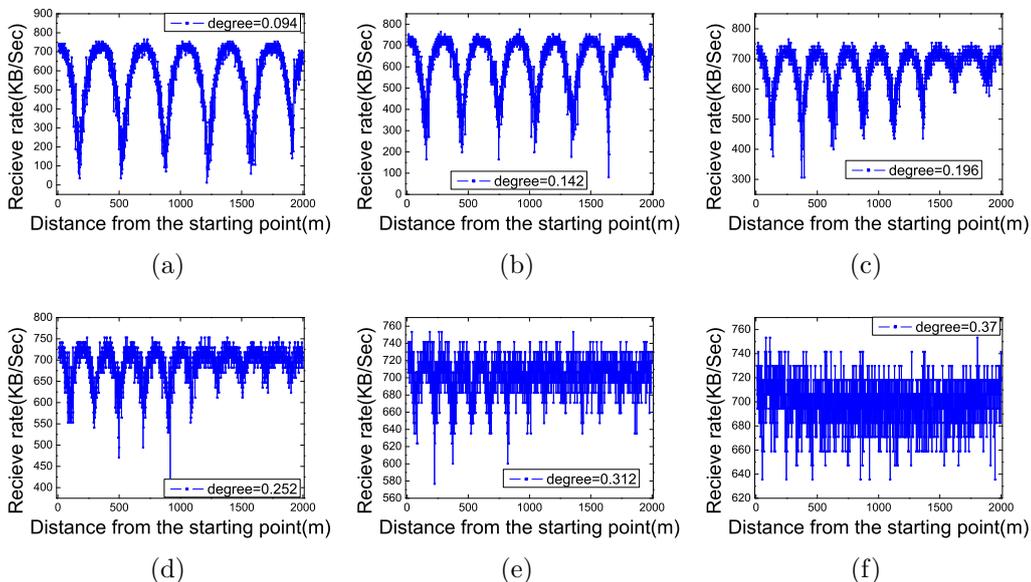
Figure 7: Received rate at MN with different degree of coverage overlapping.

7.2.1. Impact of Coverage Overlapping

We change the location of MRs in the linear network while keeping the load of MR (by fixing the number of STAs associated with the MR and also fixing the load generated by each STA) to be the same to investigate how the degree of coverage overlapping affects the mobile user's received rate. Fig.7 shows the received rate at the mobile user under our pre-scheduled handoff scheme with different degree of coverage overlapping whose values increase from Fig.7(a) to Fig.7(f). As this simulation is done in a linear network of 2000 m, the maximum distance from the starting point of the x-axis is 2000 m. Note the y-axis ranges in different Figure (7(a))-(7(f)) are different. For example, when the degree equals to 0.094 in Fig. 7(a), the data rate is within [0, 750], while in Fig. 7(f), the data rate is within [620, 750]. It is clearly shown that the mobile user moving along the path with a larger overlapping degree obtains a larger minimum received rate. In Fig.7, we observe that the received rate changes periodically and degrades in the middle of the path between two MRs. This is mainly because the strength of the received signal becomes weak when the mobile user is in the central area of the link and

26

thus farther away from the MRs. Also, the received rate reduces as a result of the handoff.

7.2.2. Impact of MR's Load

To investigate how MR's load affects the received rate, we vary the MR load by changing the data rate of the STAs associated with the MR while keeping the coverage overlapping at the same degree. We have evaluated the performance by varying MR's load under different degree of coverage overlapping, and due to the page limit, we only show results for degree $d = 0.252$ in Fig.8. All STAs associating with the same MR share and compete to access the wireless media. The actual rate of an individual mobile user depends on how the wireless channel capacity is allocated[46]. From Fig.8, we can see the higher the MR's average load is (average rate measured at MR before MN's association in Fig.8), the lower the received rate the mobile node can obtain.
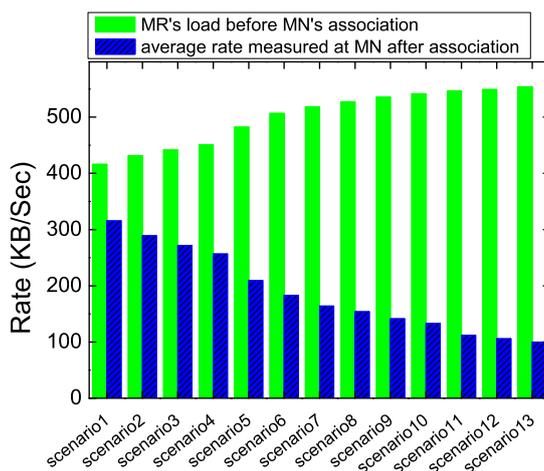
Figure 8: Received rate at MN with different MR load.

7.2.3. Impact of Handoff Strategy

To further investigate how the handoff strategy impacts the received rate, we simulate five handoff scheduling schemes with different handoff strategies. The first scheme is our pre-scheduled scheme, and the second scheme has the handoff triggered in the middle of each coverage link with the handoff procedure following the standard Wi-Fi protocol. The third scheme has

its connection procedure follow the standard Wi-Fi protocol, that is, both handoff triggering and handoff procedures (including the channel scanning and access point association) follow the standard Wi-Fi protocol. The fourth scheme is based on neighbor graph in [4, 5], where an MN only scans the working channels of its neighbor APs based on the history record. The fifth scheme follows the concept in [24] in which the next AP for the mobile user is predicted and handoff can be prepared in advance. From Fig.9, there is only a slight rate degradation in our pre-scheduled scheme, while there is a large rate degradation in other three handoff schemes due to more frequent handoffs or higher handoff delay. It is well known that the scanning delay is much larger than the association delay [6]. The received rates of the second and third schemes even reach 0 sometimes as a result of the long scanning delay. Compared to the second and third scheme, in the fourth and the fifth scheme, neighbor graph and prediction technique can be utilized for reducing the scanning delay and thus to improve the received rate.
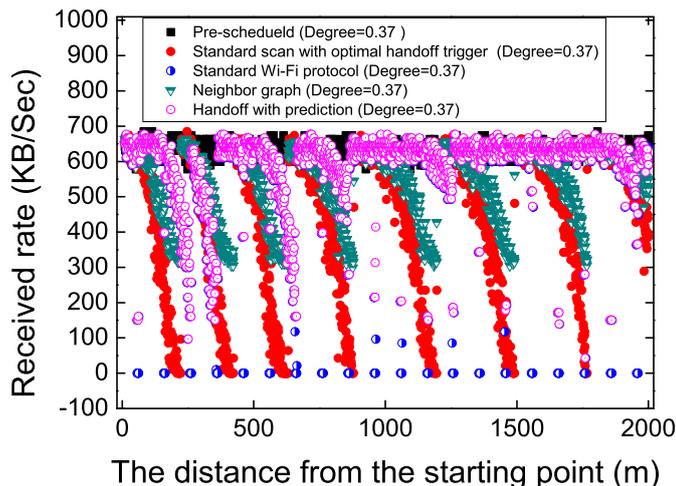
Figure 9: Received rate at MN under different handoff strategy.

A Wi-Fi covered path consists of multiple coverage links. In the current method, STA moves along a path without the knowledge of the WiFi coverage conditions, and the handoff is triggered when the signal strength is below a threshold. In our scheme, the path selection also takes into account the WiFi coverage conditions. Thus an STA can move along the links with better data rate and the next MR (say MR_B) to move to is known in advance. When the signal from MR_B is higher than a threshold, the STA is controlled to perform

28

handoff and change its association to MR_B. Therefore, mobile users can achieve higher transmission data rate. The simulation results confirm that the connection strategy proposed in our pre-scheduled handoff scheme can effectively reduce the scanning delay and improve the network performance during handoff.

7.2.4. Statistic Method to Set w_{th}

Based on numerous simulation results, we apply some statistics techniques to find the relationship between the value of contact weight and the received data rate. Contact weight is a link metric defined in Eq(1), which is calculated based on the value of the degree of coverage overlapping among MRs and the traffic load of MRs. As a mobile user may receive different rates according to his location, it makes the mapping from the contact weight to the received rate difficult. From the probability theory and statistics, the cumulative distribution function (CDF) describes the probability for a real-valued random variable X to have a value less than or equal to x, and $F(x) = p(X \leq x)$. Therefore, to better reflect the relationship between the value of contact weight and the data rate, we use CDF as the statistic tool to capture the simulation results.

Fig.10 shows the CDF results of the received rate under different contact weights. As there are many ContactWeight values, to illustrate our method, we take 7 ContactWeight values $a = 0.0002$, $b = 0.0003$, $c = 0.0004$, $d = 0.0005$, $e = 0.0006$, $f = 0.0007$, $g = 0.0008$ as examples to show the relationship. For example, if a user's requirement is that the received rate should be larger than 250KB/Sec, the w_{th} should be set to the contact weight=d; If a user's rate requirement is that the received rate should be larger than 400 KB/Sec with the probability larger than 0.8, the w_{th} should be set to the contact weight=b.

7.3. Simulation Results (Part 2)

7.3.1. Performance Comparison

As the literature routing algorithms and routing metrics are designed for finding a packet transmission route not a physical road path that a human user could walk along, there is no base for us to compare with these schemes.

In the following, five different scheduling schemes are simulated for performance comparison in the random network. The mobile users are scheduled to move along different paths under different scheduling schemes. Different mobility and connection scheduling strategies are used in the schemes.
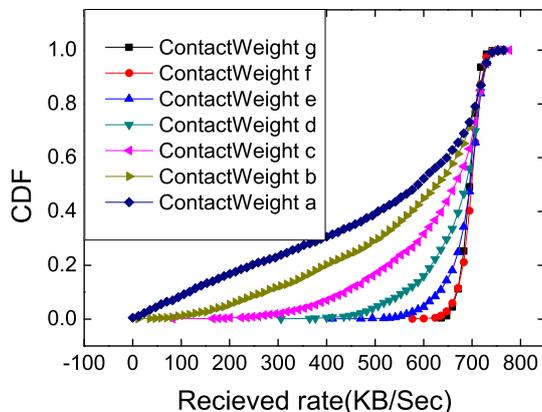
Figure 10: CDF of received rate at MN with different contact weight.

The first is our Pre-scheduled scheme with a QoS enforcement, denoted as "Pre-scheduled path1". It runs as follows. Assuming a mobile user requires the class C_i wireless service with the required rate larger than 350KB/Sec, according to Fig.10, the w_{th_i} is set to the contact weight=e for this class. Then, the optimal path of class C_i (denoted as path 1) is selected for the mobile user, and both the mobility and connection are scheduled on the path.

The second is a mobility scheduling scheme on the path 1, denoted as "M-scheduled path1", in which only the mobility is scheduled, while the connection procedure follows the standard Wi-Fi protocol.

The third is the pre-scheduled scheme with the minimum handoff frequency but without considering the QoS requirement, denoted as "M-C-scheduled path2", in which the path with the minimum handoff count is selected (path 2), and both the mobility and connection are scheduled for the mobile user on the path.

The fourth is the mobility scheduling scheme on the path 2, denoted as "M-scheduled path2", in which only the mobility is scheduled on the path.

The final one is the mobility scheduling scheme with the minimum length of path, denoted as "M-scheduled path3", in which the path with the minimum physical distance is selected (path 3), and only the mobility is scheduled on the path.

The CDF curves of the received rates at MN under different scheduling schemes are shown in Fig.11. More statistical results, including the mean

received rate, the standard deviation of the received rate, and the minimum and maximum received rate, handoff count, packet loss ratio (the ratio of the number of lost packets to the total number of packets transmitted), None-WiFi length ratio (the ratio of the length of path segment without WiFi covered to the length of the whole path)are listed in Table 1.

Our pre-scheduled handoff scheme provides the largest mean received rate with the minimum standard deviation, which proves our scheme can achieve a high performance and a high stability, and outperforms other schemes. The minimum received rate under our pre-scheduled scheme is 375.41 KB/Sec, and satisfies the required rate 350 KB/Sec. Because the contact weights of different links in the path are different, the CDF of the received rate in Fig.11 is different from the CDF of the received rate of ContactWeight=e in Fig.10.
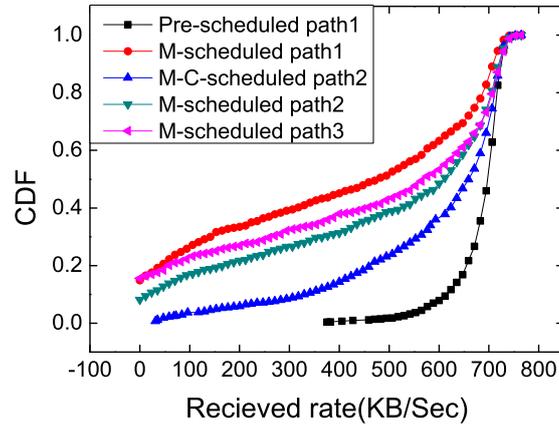
Figure 11: CDF of received rate at MN under different scheduling schemes.

Table 1: Performance comparison

Scheduling scheme	Mean Rate	Standard Deviation	Minimum Rate	Maximum Rate	Handoff Count	Packet Loss Ratio	None-WiFi Length Ratio
Pre-scheduled path1	686.49439	45.36459	375.41	765.05	9	0	0
M-scheduled path1	400.22941	233.16399	0	753.28	8	1%	0
M-C-scheduled path2	590.1108	175.32165	23.54	765.05	7	0	0
M-scheduled path2	496.26707	237.97418	0	765.05	7	1%	0
M-scheduled path3	452.25122	267.04595	0	765.05	12	3%	30%

The continuous Internet access can be obtained in M-C-scheduled path2 with the minimum received rate of 23.54KB/Sec. Because this scheme as-

sumes the optimal scheduling strategies for mobility and connection and selects the path 2 based on the coverage connection considered in this paper, the intermittent Wi-Fi coverage is avoided. However, there is no minimum rate guarantee. The minimum received rate in M-C-scheduled path 2 is only about 6 percent of the minimum received rate in Pre-scheduled path 1.

Although a mobile user moves on a fully Wi-Fi covered path, the mobile user cannot obtain the continuous Seamless Internet access in M-scheduled path1 and M-scheduled path2. Similar to the results of the linear network in Fig.9, as the connection procedure is not optimized in these two schemes, a large handoff delay leads to the outage of transmission and thus the packet loss.

In Pre-scheduled path1 and M-C-scheduled path2, the connection is well scheduled to significantly reduce the scanning delay, so there is no packet loss in these two schemes. Although M-scheduled path3, M-scheduled path1, and M-scheduled path2 adopt the same connection procedure, the packet loss ratio in M-scheduled path3 is the largest, as it simply takes the shortest path without any Wi-Fi covered guarantee. When mobile users move to the coverage boundary of MRs, the weak signal leads to higher communication outage rate and thus large packet loss ratio.

Among all the five schemes, M-C-scheduled path2 and M-scheduled path2 have the lowest handoff count as path2 is chosen with the minimum handoff frequency. However, path2 is chosen without considering the QoS requirement. Although in M-C-scheduled path2 a user moves along the path with the minimum handoff count, its received rate is much worse than that under Pre-scheduled path1.

7.3.2. Performance Comparison With Different STA Distribution

To study the performance of a network with non-uniform STA distribution, Fig.12 shows the performance result under two non-uniform STA distribution scenarios. We observe that, although our proposed contact weight definition may either overestimate or underestimate the coverage overlap of two neighboring MRs, the minimum received rate achieved under our scheme is still much better than that achieved under M-scheduled path3 in which the mobile user moves along the shortest path. These results prove the effectiveness of our scheme in finding a good WiFi covered path.

In summary, all the results show that our pre-scheduled handoff scheme can provide appealing high data rates with the minimum handoff delay for mobile users.
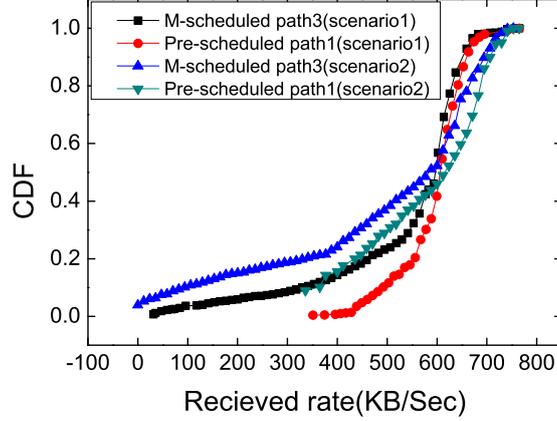
Figure 12: Performance Comparison With Different STA Distribution.

8. Algorithm Overhead

To enable the proposed routing scheme, it is important to obtain the metric ContactWeight. We analyze the message cost to obtain the necessary parameters and the cost for calculating the metric.

The Bloom filter vector that represents the STA set covered by MRs needs to be exchanged among MRs. As discussed earlier, one major reason for us to choose the Bloom filter is that it is a succinct data structure and is widely applied in the distributed system to enable low-cost message exchanges.

Given SBF_1 and SBF_2 which represent the two STA sets covered by MR_1 and MR_2 respectively, to calculate the degree of coverage overlapping of a pair of MRs as shown in Eq(2), we first make use of the union property of Bloom filter in Theorem 1 to estimate the cardinality of the set union using the Theorem 2. We then find the degree of coverage overlapping based on the estimated set sizes of S_1, S_2, $S_1 \cup S_2$, and $S_1 \cap S_2$.

The cost on calculating the degree of coverage overlapping is low. Given a Bloom filter, to estimate the size of STA set represented by the Bloom filter, the cost for counting the number of "1" in a Bloom filter of length m is $O(m)$. Also, our proposed probabilistic estimation is performed based on simple binary operations of Bloom filter strings, which can be easily implemented on well-designed hardware at the low complexity.

The load of an MR can be estimated based on the total traffic of its associated users, which can be easily obtained from the existing 802.11 MR

33

products. After each MR estimates its load information, the information will be exchanged among its neighboring MRs to facilitate the calculation of the ContactWeight of a link. The message exchange cost is low.

For the message cost in path finding procedure, by defining the link cost in (7), the original path finding problem in (6) is transformed into the problem in (8) to look for a path with the minimum accumulated cost of all links. To solve problem, only the information of link cost (defined in (7)) needs to be exchanged among nodes to calculate the path table. The communication cost is not high. Moreover, although the handoff count is considered in the path finding, this can be automatically taken care when minimizing the path cost (8) without incurring extra cost.

Therefore, our pre-handoff scheme introduces simple calculation to measure the proposed metric. The communciation cost and calculation cost introduced by our pre-handoff scheme is low.

9. CONCLUSION

In this paper, we study a novel service-aware and seamless Internet access problem, and propose a fully distributed and pre-scheduled handoff scheme to solve the problem. In our scheme, mobility and connection are well scheduled for the mobile users to continuously access Internet with his expected data rate. Several novel techniques are proposed. First, we propose two novel path finding metrics (*contact weight* and *handoff count*) which are closely related to the quality of data service that a mobile user might experience when moving through the network. Second, to measure the path metric in a practical mesh network with irregular and dynamic MR coverage, we first propose a light-weight polling algorithm which exploits Bloom filter to quickly discover the STAs covered in MR, and then further apply the Bloom filter to estimate the degree of coverage overlapping among MRs. Third, based on the metrics, to flexibly support performance requirements of mobile users, we propose a fully distributed multi-service-class path finding algorithm. Finally, we propose a cross-layer pre-scheduled connection strategy where both L_2 and L_3 handoffs are pre-scheduled on the path for the mobile users. The performance results demonstrate that our pre-scheduled handoff scheme can effectively support network connectivity and with a much higher data rate, as compared with the reference schemes.

Instead of relying on the current opportunistic handoff schemes, providing a better-covered path may allow the support of some advanced wireless

applications which have higher requirements on network connectivity and transmission quality. For example, a user may hold a wireless terminal to monitor a remote site. With the support of the pre-scheduled handoff, the user can obtain a seamless network access and high quality surveillance video. This application can be used by policeman on the move to better monitor a public space for security.

In this paper, we focus on finding a Wi-Fi covered path to provide the mobile user seamless and high received date rate. To provide the mobile users with flexible QoS support, in our future work, we will take account the delay and jitter into path finding procedure.

Acknowledgment

[1] S. Vural, D. Wei, K. Moessner, Survey of experimental evaluation studies for wireless mesh network deployments in urban areas towards ubiquitous internet, IEEE Communications Surveys Tutorials 15 (1) (2013) 223–239. doi:10.1109/SURV.2012.021312.00018.

[2] P. Guo, J. Wang, X. H. Geng, C. S. Kim, J.-U. Kim, A variable threshold-value authentication architecture for wireless mesh networks, Journal of Internet Technology 15 (6) (2014) 929–935.

[3] J. Cao, K. Xie, W. Wu, C. Liu, G. Yao, W. Feng, Y. Zou, J. Wen, C. Zhang, X. Xiao, et al., Hawk: Real-world implementation of high-performance heterogeneous wireless network for internet access, in: ICDCS Workshop on Distributed Computing Systems, IEEE, 2009.

[4] A. Mishra, M. Shin, W. A. Arbaush, Context caching using neighbor graphs for fast handoffs in a wireless network, in: INFOCOM 2004.

[5] S. Pack, H. Jung, T. Kwon, Y. Choi, SNC: a selective neighbor caching scheme for fast handoff in IEEE 802.11 wireless networks, ACM SIGMOBILE Mobile Computing and Communications Review 9 (4) (2005) 39–49.

[6] S. Pack, J. Choi, T. Kwon, Y. Choi, Fast-handoff support in IEEE 802.11 wireless networks, IEEE Communications Surveys & Tutorials 9 (1) (2007) 2–12.

[7] H. Soliman, L. Bellier, K. E. Malki, Hierarchical mobile IPv6 mobility management (HMIPv6).

[8] R. Koodli, Fast handovers for mobile IPv6, IETF RFC 4068.

[9] B. Rong, Y. Qian, K. Lu, R. Hu, M. Kadoch, Mobile-agent-based hand-off in wireless mesh networks: Architecture and call admission control, IEEE Transactions on Vehicular Technology 58 (8) (2009) 4565–4575. doi:10.1109/TVT.2009.2021062.

[10] M. Buddhikot, A. Hari, K. Singh, S. Miller, Mobilenat: A new technique for mobility across heterogeneous address spaces, Mobile Networks and Applications 10 (3) (2005) 289–302.

[11] K. Ramachandran, M. Buddhikot, G. Chandranmenon, S. Miller, E. Belding-Royer, K. Almeroth, On the design and implementation of infrastructure mesh networks, in: WiMesh 2005.

[12] S. Speicher, C. H. Cap, Fast layer 3 handoffs in AODV-based IEEE 802.11 wireless mesh networks, in: ISWCS 2006.

[13] S. Speicher, OLSR-FastSync: fast post-handoff route discovery in wireless mesh networks, in: VTC 2006.

[14] R. Huang, C. Zhang, Y. Fang, A mobility management scheme for wireless mesh networks, in: GLOBECOM 2007.

[15] V. Navda, A. Kashyap, S. R. Das, Design and evaluation of imesh: an infrastructure-mode wireless mesh network, in: WoWMoM 2005.

[16] Y. Amir, C. Danilov, M. Hilsdale, R. Musloiu-Elefteri, N. Rivera, Fast handoff for seamless wireless mesh networks, in: Proceedings of the 4th international conference on Mobile systems, applications and services, ACM, 2006.

[17] Y. Zhang, Handoff performance in wireless mobile networks with unreliable fading channel, IEEE Transactions on Mobile Computing 9 (2) (2010) 188–200.

[18] R. Wishart, M. Portmann, J. Indulska, Evaluation of wireless mesh network handoff approaches for public safety and disaster recovery networks, in: ATNAC 2008.

[19] Y. Yang, D. Zhao, Joint handoff and energy management for a wireless mesh network, in: ICC 2011.

[20] H. Li, J. Xie, An adaptive channel scheduling design for multi-hop handoff delay reduction in internet-based wireless mesh networks, IEEE Transactions on Wireless Communications 12 (10) (2013) 5258–5270.

[21] H. Li, J. Xie, X. Wang, A channel splitting strategy for reducing handoff delay in internet-based wireless mesh networks, IEEE Transactions on Vehicular Technology 61 (6) (2012) 2740–2752. doi:10.1109/TVT.2012.2195512.

[22] W. Zhao, J. Xie, IMeX: Intergateway cross-layer handoffs in internet-based infrastructure wireless mesh networks, IEEE Transactions on Mobile Computing 11 (10) (2012) 1585–1600. doi:10.1109/TMC.2011.192.

[23] P. Deshpande, A. Kashyap, C. Sung, S. R. Das, Predictive methods for improved vehicular WiFi access, in: MobiSys 2009.

[24] Y. Chon, H. Shin, E. Talipov, H. Cha, Evaluating mobility models for temporal prediction with high-granularity mobility data, in: PerCom 2012.

[25] N. Vishnu, S. Anand Prabhu, D. Kannan, T.-G. Andreas, D. Samir, MobiSteer: using steerable beam directional antenna for vehicular network access, in: MobiSys 2007.

[26] A. P. Subramanian, P. Deshpande, G. Jie, S. R. Das, Drive-by localization of roadside WiFi networks, in: INFOCOM 2008.

[27] V. Bychkovsky, B. Hull, A. Miu, H. Balakrishnan, S. Madden, A measurement study of vehicular internet access using in situ wi-fi networks, in: MobiCom 2006.

[28] J. Ott, D. Kutscher, A disconnection-tolerant transport for drive-thru internet environments, in: INFOCOM 2005.

[29] J. Ott, D. Kutscher, Drive-thru internet: Ieee 802.11 b for" automobile" users, in: INFOCOM 2004.

[30] V. Navda, A. P. Subramanian, K. Dhanasekaran, A. Timm-Giel, S. Das, Mobisteer: using steerable beam directional antenna for vehicular network access, in: Proceedings of the 5th international conference on Mobile systems, applications and services, ACM, 2007, pp. 192–205.

[31] A. Mishra, M. Shin, W. Arbaugh, An empirical analysis of the IEEE 802.11 MAC layer handoff process, ACM SIGCOMM Computer Communication Review 33 (2) (2003) 93–102.

[32] H. Li, J. Xie, GaS: A gateway scheduling-based handoff scheme in single-radio infrastructure wireless mesh networks, in: INFOCOM, 2013.

[33] W. Zhao, J. Xie, A novel xcast-based caching architecture for inter-gateway handoffs in infrastructure wireless mesh networks, in: INFOCOM, 2010.

[34] V. Navda, S. R. Das, Design and evaluation of iMesh: An infrastructure-mode wireless mesh network, in: WoWMoM 2005.

[35] R. K. Sheshadri, D. Koutsonikolas, An experimental study of routing metrics in 802.11 n wireless mesh networks, IEEE Transactions on Mobile Computing 13 (12) (2014) 2719–2733.

[36] D. Couto, D. Aguayo, J. Bicket, R. Morris, a high-throughput path metric for multi-hop wireless routing, Wireless Networks 11 (2005) 419–434. doi:10.1007/s11276-005-1766-z.
URL http://dx.doi.org/10.1007/s11276-005-1766-z

[37] R. Draves, J. Padhye, B. Zill, Routing in multi-radio, multi-hop wireless mesh networks, in: MobiCom, 2004.

[38] Y. Yang, J. Wang, R. Kravets, Load-balanced routing for mesh networks, ACM SIGMOBILE Mobile Computing and Communications Review 10 (4) (2006) 3–5.

[39] H. Wu, F. Yang, K. Tan, J. Chen, Q. Zhang, Z. Zhang, Distributed channel assignment and routing in multiradio multichannel multihop wireless networks, IEEE Journal on Selected Areas in Communications 24 (11) (2006) 1972 –1983. doi:10.1109/JSAC.2006.881638.

[40] K. Xie, X. Wang, X. Liu, J. Wen, J. Cao, Interference-aware cooperative communication in multi-radio multi-channel wireless networks, IEEE Transactions on Computers 65 (5) (2016) 1528–1542.

[41] K. Xie, X. Wang, J. Wen, J. Cao, Cooperative routing with relay assignment in multiradio multihop wireless networks, IEEE/ACM Transactions on Networking 24 (2) (2016) 859–872.

[42] A. Broder, M. Mitzenmacher, Network applications of bloom filters: A survey, Internet Mathematics (2002) 636–646.

[43] T. Hagerup, C. Rüb, A guided tour of chernoff bounds, Information processing letters 33 (6) (1990) 305–308.

[44] D. Bertsekas, R. Gallager, Data networks. 2nd ed., in: Prentice Hall 1991.

[45] S.-Y. Wang, C.-C. Lin, NCTUns 6.0: A simulator for advanced wireless vehicular network research, in: VTC 2010, 2010.

[46] G. Tan, J. Guttag, Time-based fairness improves performance in multi-rate Wlans, in: USENIX Annual Technical Conference, 2004.