Hsieh and Kao *EURASIP Journal on Wireless Communications and Networking* 2011, **2011**:30
http://jwcn.eurasipjournals.com/content/2011/1/30

EURASIP Journal on Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH Open Access

Handoff optimization in 802.11 wireless networks

IP Hsieh and Shang-Juh Kao*

Abstract

In 802.11 wireless networks, a complete handoff procedure for a mobile node requires access point (AP) selection, AP switch, call admission control (CAC), IP address re-allocation, and network re-configuration. Most current handoff schemes deal only with either AP selection or IP address re-allocation. In this paper, an integrated handoff procedure is proposed. First, AP selection is accomplished by choosing an AP with the lowest channel utilization and smaller number of associated users. The information about load of each AP is reported through modified beacon frames. In the case of adopting load-based AP selection, the average throughput can be increased up to 56%, as opposed to pure SNR-based AP selection. Next, both CAC and IP address pre-fetch are performed simultaneously through the simplified DHCP procedure. Specifically, efficient limited fractional guard channel policy (ELFGCP) is proposed for the CAC phase. By adopting ELFGCP, the failure probability can be reduced as much as 45% from conventional LFGCP. Finally, the simulation results demonstrate the applicability of the integrated approach, and the overall disconnection time due to handoff can be reduced from 2.9 to 0.004 s using traditional handoff procedures.

Keywords: AP selection, Handoff, Call admission control, Efficient LFGCP

1. Introduction

With the rapid growth of the deployment of WiFi [1] devices, mobile users can easily access Internet resources. Since handoff is indispensable to a mobile user changing location, it is imperative to achieve seamless handoff so that the connection is maintained. Seamless handoff requires minimizing the time to process handoff and the rate of packet loss. To address packet loss, several schemes [2,3] have been proposed that the corresponding node duplicates the data packets among relevant APs. However, these duplicated packets may overload network devices. Therefore, a better approach is to reduce the time spent in handoffs. Handoffs occur across several layers. For example, the handoff in the network layer deals with network address re-allocation, whereas the data link layer determines a target access point (AP). In order to reduce the time in processing a handoff, its operations must be integrated across layers. Even though lower-layer protocols involved in handoff are simplex, integration with various application layer protocols is not trivial. Even though an application layer handoff approach, proactive and adaptive handover system [4], was proposed using session initiation protocol (SIP) [5], some network and data link layer operations can be integrated, for example an IP address pre-fetching and link layer call admission.

The conventional handoff process in data-link and network layers includes five phases: AP selection, AP switch, call admission control (CAC), IP address re-allocation, and network re-configuration, as shown in Figure 1. These phases are activated only after the mobile node (MN) has disconnected from its associated AP. In the AP selection phase, the MN evaluates all nearby APs, by using an indicator such as signal strength. Next, the MN deals with the AP association in the AP switch phase. An *ASSOCIATION RESPONSE* will be received by the MN once the *ASSOCIATION REQUEST* is granted in the CAC phase, which allows the MN to then access the network following by the IP address re-allocation phase. Then the MN retrieves a new IP address via DHCP. Finally, the network configuration is changed accordingly and the MN operates as usual.

In order to minimize handoff duration, several operations could be dealt with before disconnect. Since IP address re-allocation is the most time-consuming phase in the conventional handoff process [6], several IP address pre-fetch schemes, which are carried out via DHCP Relay [7] or BOOTP [8,9], have been proposed

* Correspondence: sjkao@cs.nchu.edu.tw
Department of Computer Science and Engineering, National Chung-Hsing University, 250 Kuo-Kuang Rd., Taichung, 40227, Taiwan

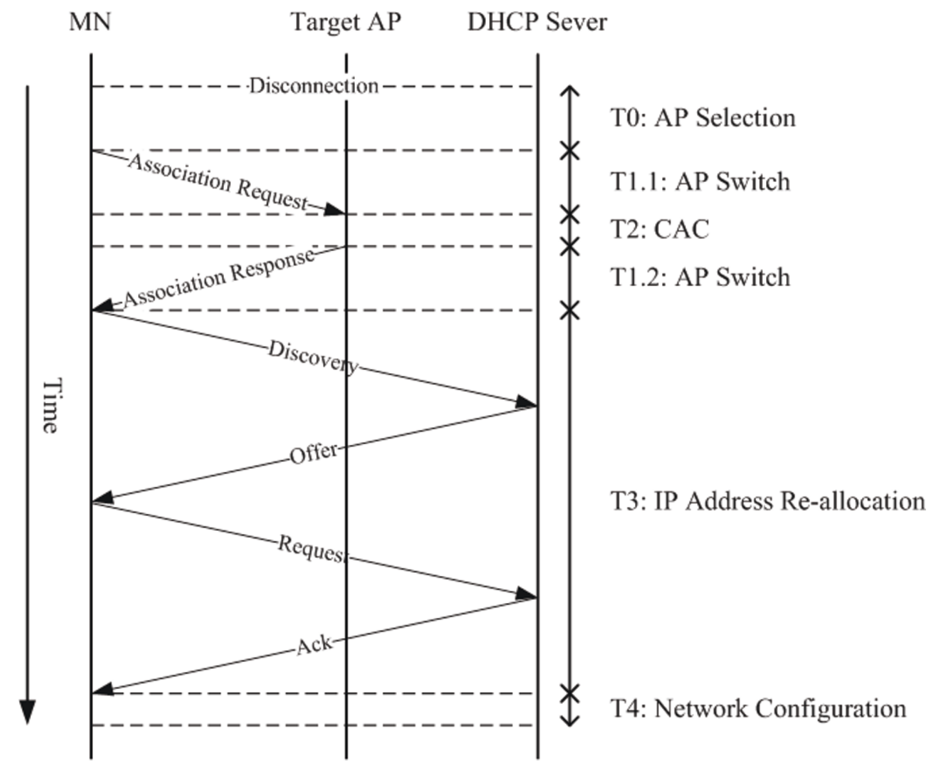

Figure 1 Conventional handoff process.

[3,10]. A target network should be determined before an IP address is obtained, hence the AP selection phase has to be accomplished before disconnect. The work flow of IP address pre-fetch is shown in Figure 2. The advantage of network address pre-fetch is to reduce the time until disconnect, during which a valid IP address is granted.

Although considerable time can be saved through IP address pre-fetch, one potential problem is that the MN may not be granted to access to the target network after the disassociation. If the access request is denied by the target AP during the CAC phase, the traditional handoff process will be resumed. This leads to greater time consumption, represented by T1 and T2 in Figure 2, than does the traditional handoff procedure. In this paper, we propose to further move the CAC phase before physical disconnect. Thus, the handoff latency will be reduced to MAX(T1, T4), as shown in Figure 3. AP switch phase (T1) is the duration from the MN disassociating with the serving AP to its associating with the target AP. Network configuration phase (T4) is the time for the MN to set up the newly retrieved IP address and other network parameters, such as network mask and IP addresses of DNS servers. Since T1 and T4 are indispensable but can be simultaneously operated after disassociation, the duration of disconnect is minimized. In contrast to the traditional handoff process, all other procedures of AP selection, CAC, and IP address pre-fetch are finished before disconnect from the serving AP.

A new AP selection mechanism, a simplified DHCP procedure, and an efficient CAC algorithm are proposed in this paper to achieve the handoff optimization. In the AP selection phase, the information about user count and channel utilization is included in the broadcast beacon frame. Thus, T0 can be reduced, and MNs is able to associate with the AP of lower user load. And, the transmission throughput is increased. Since the conventional DHCP procedure is time-consumed, the standard message flow is modified as shown in Figure 3. Once the MN receives the *OFFER* message, the MN accepts the assigned IP address and lease unconditionally. Therefore, T3 is reduced significantly. Since the execution duration of the above two phases is shortened, the success probability of both phases is raised before the physical disconnection. In the proposed CAC algorithm, resources allocation is reconfigured depending on the ratio of handoff calls to new calls. Though higher priority is assigned to handoff MNs, the resources are still allocated to new coming MNs whenever possible. Thus,

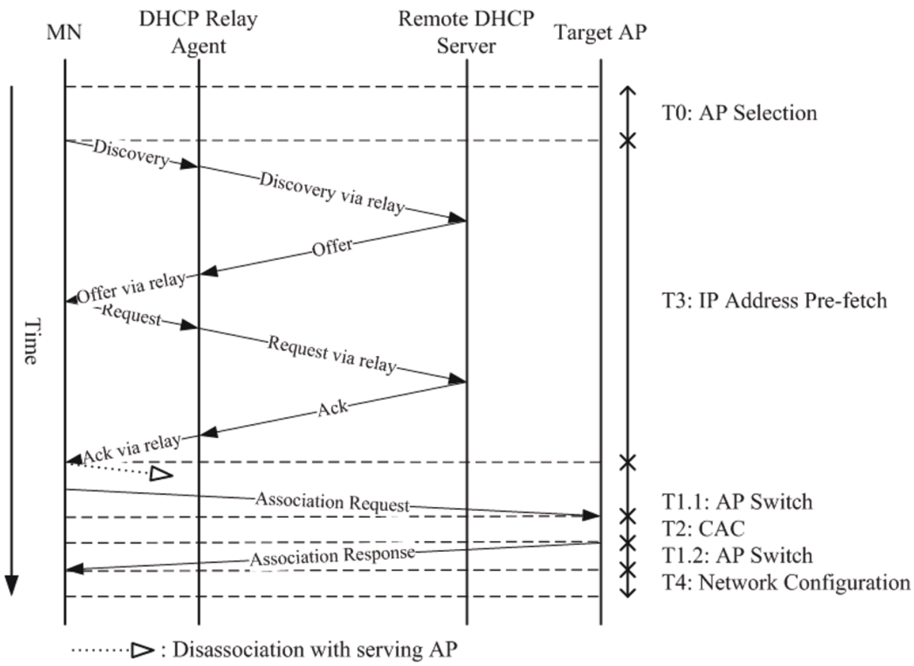

Figure 2 Handoff process with IP address pre-fetch.

the possibility for MNs to successfully retrieve the grant is improved.

The remainder of this paper is organized as follows. Related work on handoff schemes is reviewed. Then, an optimized fast handoff (FHO) procedure is proposed. The system architecture and corresponding algorithm are presented. Following the system description, a simulation system using OMNET++ and the experimental results are illustrated. Finally, conclusions and future work are given.

2.1. Related studies on AP selection

Performance indicators are essential to optimizing AP selection. In this section, various indicators for different network layers of most current AP selection schemes are reviewed.

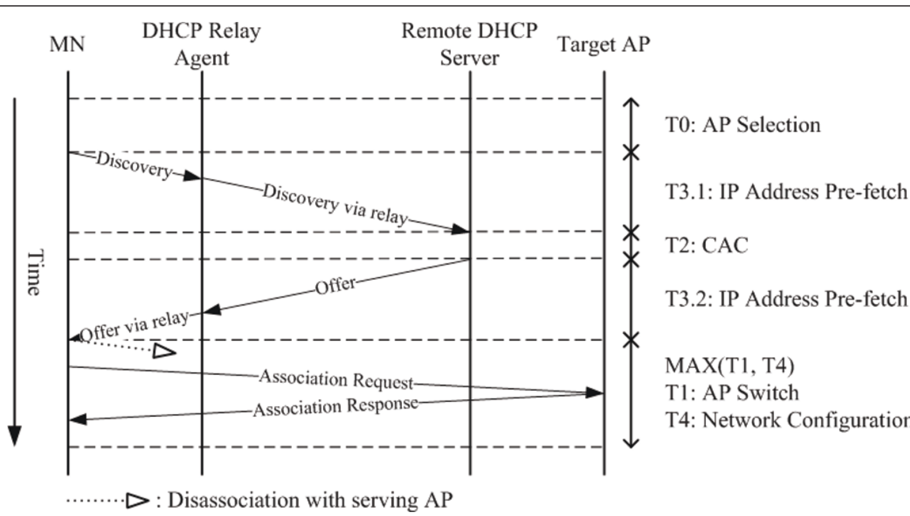

Figure 3 The proposed handoff process.

2.1.1. Physical layer indicators

At the physical layer, signal-to-noise ratio (SNR), channel influence [11] and channel utilization [1], are the most commonly used performance indicators. SNR can be detected directly from the MNs by listening or scanning reachable APs. When this indicator is adopted by itself, the MNs will simply associate with the AP having the strongest SNR. Although this strategy is easy to implement, it may cause MNs to associate primarily with a few dedicated APs, which may lead to load imbalance [12-14]. Channel interference may also be used as a physical layer indicator, as an MN attempts to associate with the AP whose channel is less interfered. The channel utilization, which has been defined in the standard of IEEE 802.11-2007 [1] as a resource usage indicator, could be detected by APs and MNs. It is defined as the percentage of time that an AP has sensed the channel being busy. Therefore, this indicator can also be taken as a loading factor of APs. If the indicator is detected by MNs via listen the media, the value would be incorrect due to the hidden node problem [15].

2.1.2. Data-link layer indicators

At the data-link layer, frame error rate [13], transmission rate [16-18], and contention level [19] are usually taken as the performance indicators for AP selection. Due to the popularity of multimedia services, QoS is becoming critical. QoS typically refers to low frame error rate or high transmission rate at the data-link layer. Both of these can be calculated by transmitting frames or by retrieving objects in a related management information base via simple network management protocol [20]. The disadvantage of this QoS measure is the huge time overhead in data collection. In [19], mean probe delay (MPD) was proposed and used as an indicator of an AP's load. The MN delivers four probe requests in sequence and measures the elapsed time of the respective responses. The AP with the lowest time delay is selected.

2.1.3. Indicators above the data-link layer

Above the data-link layer, throughput [6] and QoS [15,21] are the common indicators for AP selection. The problem of gathering high-layer indicators lies in the time latency of AP association, which has to be established before starting the measurement. In contrast to the indicators used in lower layers, such as SNR and MPD, the higher-layer AP selection indicators are not easily derived directly. The overhead of associating with the APs for obtaining the measure leads to the impracticability of higher-layer indicator measurement. Consequently, most current AP selection schemes rely on lower-layer indicators.

2.1.4. Other indicators

Aside from layer-based indicators, valuable indicators for AP selection include user count [22], potential hidden nodes [15], moving direction [23], link cost, services provision, and access technology. The easiest AP selection scheme for load balancing might be to evenly distribute application users among reachable APs. Although application-layer load distribution protocol [15] can be used to access the user count, it is an application layer protocol. This means that the overhead for retrieving user count is expensive.

Another interesting indicator is the number of hidden nodes, which are the MNs mistakenly treated as out of radio coverage. The hidden node problem [15] may cause increase in collision probability. Potential hidden terminal effect (PHTE) [15] tries to address the hidden nodes problem without *RTS* [24] and *CTS* [24] messages. Accordingly, MNs will associate with the AP with the lowest PHTE to reduce collision. The main drawback is that the extended functions of 802.11e [15] have not been implemented in most APs. Taking into consideration, the direction of an MN should effectively reduce handoff times. In order to detect the direction of an MN and discover APs that are in front of the MN, additional information, such as geographic information, is required. In the future, an MN will be capable of accessing resources via different wireless access technologies, such as WLAN, WiMAX [25], and 4G [26]. Due to differences in their specifications, the access technology should be considered accordingly. All AP selection indicators with the collecting methods, advantages, and disadvantages are summarized in Table 1.

2.2. Related studies on IP address pre-fetch

The critical problem in the IP address pre-fetch phase is how to obtain the IP address of the DHCP server in the target network. In the proposed CF-SIP [10], the MAC address of the gateway of the target network is encapsulated into the beacon frames. Thus, any MN can retrieve the IP address of the target DHCP server via BOOTP *REQUEST* and *REPLY*. Also, an IP address can be successfully assigned to an MN via a DHCP relay agent once the IP address of the DHCP server is retrieved. The complete procedure for CF-SIP is shown in Figure 4.

The major disadvantage of CF-SIP is the long processing time. It involves exchanging six messages: *BOOTP REQUEST*, *BOOTP REPLY*, *DHCP DISCOVERY*, *DHCP OFFER*, *DHCP REQUEST*, and *DHCP ACK*. Thus, three round-trip times (RTTs) are required. Long processing time may preclude the MN from successfully obtaining an IP address before disconnect from the serving AP.

Table 1 Current AP selection schemes

Indicator	Layer	Gathering method	Advantage	Disadvantage
SNR	Physical	Listen or scan by MN	Fast, stable connection; easy to implement	Load imbalance
Channel influence	Physical	Listen or scan by MN	Reduce influence	Load imbalance
Channel utilization	Physical	By APs	Load balancing accurate	Hardware support
Channel utilization	Physical	By MNs	Load balancing	Additional time Inaccurate
Frame Error Rate	Data-Link	By MNs or APs	Faster than application layer QoS estimation	Inaccurate
Transmission rate	Data-Link	By MNs or APs	Faster than application layer QoS estimation	Inaccurate
MPD	Data-Link	By MNs or APs	Load balancing	Additional time
Throughput	Network	By MNs or APs	Accurate	Additional time
User count	N/A	By APs	Load balancing	Additional time
PHTE	N/A	By MNs	Collision prevention	Additional time Hardware support
Moving direction	N/A	Bby MNs	Appropriate AP choice	Extra data collection
Link cost	N/A	User assisted	Reduce cost	Extra data collection
Service provision	N/A	User assisted	Service-based support	Extra data collection
Access technology	N/A	By MNs	Support high speed movement	Additional time

2.3. Related studies of CAC

The main reason for using CAC is to guarantee QoS parameters such as signal quality, call dropping probability (DP), packet-level parameters, and transmission rate [27]. In an 802.11 WLAN, APs are incapable of maintaining signal quality due to the movement of MNs and various interferences. Furthermore, WiFi is a contention-oriented wireless access technology. Thus, to guarantee packet-level parameters and transmission rate is almost impossible for current APs. Consequently, call DP is the only QoS parameter that could be regulated by current APs. Since users perceive dropping an ongoing call as more disruptive than blocking a new call, CAC is employed to reduce the handoff call DP. CAC could be implemented by reserving resources for handoff calls. The admission criterion can be the user

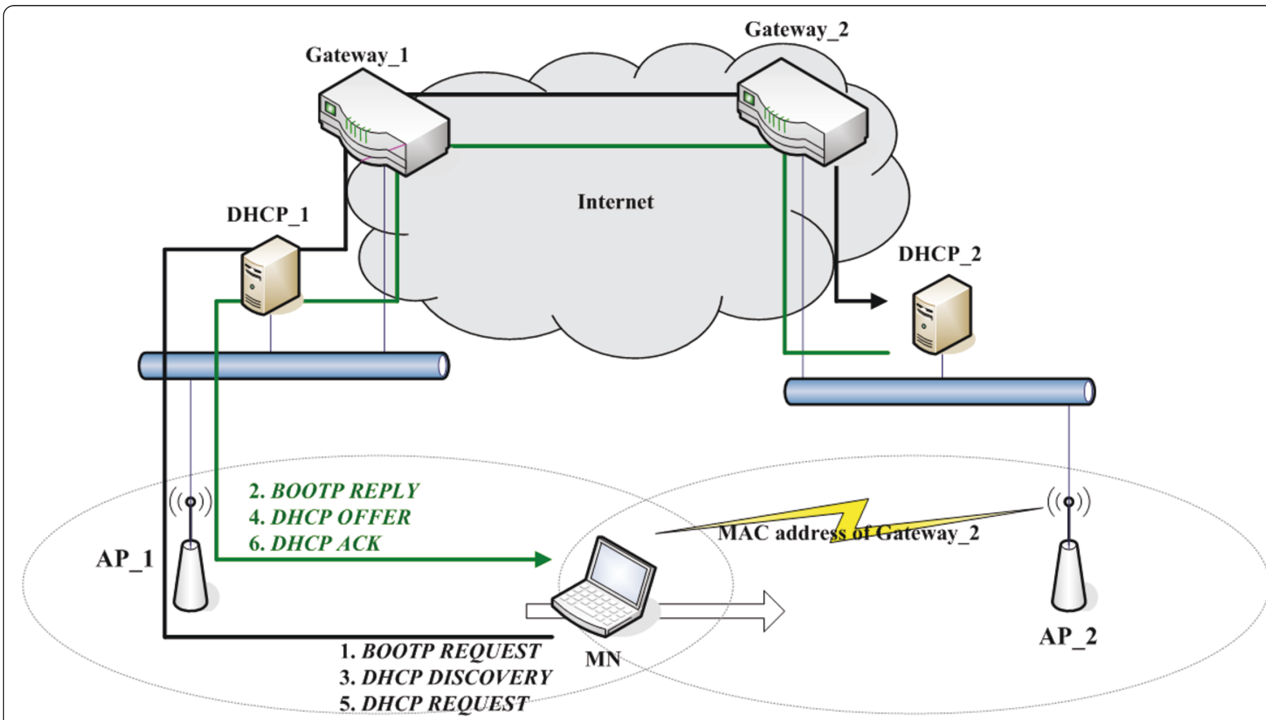

Figure 4 Work flow of CF-SIP scheme.

count, resource availability, or an estimated DP. Whatever the admission criterion, handoff calls receive less stringent admission conditions than does a new call, which might increase new call blocking probability (BP). The following subsections investigate three current CAC policies [28]: guard channel policy (GCP), fractional guard channel policy (FGCP), and limited fractional guard channel policy (LFGCP).

2.3.1. Guard channel policy

In the GCP algorithm, C is the maximum user count for an AP, and T is the lower bound of available capacity for handoff calls. When a new call arrives, it is accepted if the available capacity is $<T$. Thus, DP could be reduced significantly with large T. The algorithm of GCP is as follows:

Algorithm?

```
1 if (a handoff call arrives)
2   if (Num_Of_Occupied_Channels <C)
3      accept the call
4   else
5      reject the call
6 if (a new call arrives)
7   if (Num_Of_Occupied_Channels < (C-T))
8      accept the call
9   else
10     reject the call
```

2.3.2. Fractional GCP

In the GCP algorithm, the reserved capacity would be wasted if there are far fewer handoff calls than new calls, while the available capacity is $<T$. The BP will increase unacceptably. Thus, the FGCP algorithm dopts an external function β to handle the difficulty for new calls to be accepted. The condition random (0,1) < β (Num_Of_Occupied_Channels) is substituted for the original condition in line 7 of the GCP algorithm. This function could be defined by network administrators. For example, the function could return the reciprocal of occupied capacity.

2.3.3. Limited fractional guard channel policy

If new calls are accepted while the available capacity is low, the BP would be unacceptable. Unfortunately, new calls might be rejected even if there is sufficient capacity. These problems can be avoided by integrating the LFGCP algorithm with the above two CAC algorithms. While the available capacities are larger than T, GCP is adopted; otherwise, FGCP is adopted. The LFGCP algorithm replaces the pseudocode for dealing with new calls with the following:

```
6 if (a new call arrives)
7   if (Num_Of_Occupied_Channels <T)
8      accept the call
9   else if (Num_Of_Occupied_Channels < (C - T))
&& (random (0,1) < β (Num_Of_Occupied_Channels))
10     accept the call
11  else
12     reject the call
```

3. System architecture

The essential requirement of a FHO process is short handoff delay. As shown in Figure 3, the minimum handoff delay is MAX(T1, T4), and all other phases could be done in advance. Streamlining these pre-procedures is critical to the success of FHO. In this paper, an architecture of FHO system that integrates AP selection, CAC, and IP address pre-fetch is proposed.

3.1. AP architecture

In the proposed handoff process, several special functions, such as DHCP relay and CAC, must be integrated into APs. The proposed AP architecture is shown in Figure 5.

The communication module is responsible to deal with both wireless and wired messages, such as broadcasting *Beacons*, receiving *ASSOCIATION REQUESTs*, and sending *ASSOCIATION RESPONSEs*. The DCHP Module is not only a DHCP server but also a DHCP relay agent. The module is capable of relaying *DHCP DISCOVERY* to the target AP and relaying *DHCP OFFER* to mobile users. A given MN association is approved by the CAC Module according to the efficient limited fractional guard channel policy (ELFGCP) algorithm. All information, such as the assigned IP addresses and the MAC addresses of all permitted MNs, or recorded in the info store (IS).

3.2. AP selection

An efficient AP selection scheme involves gathering and analyzing performance indicators to establish an MN association. Which indicator should be collected depends on the purpose of the user's application. For instance, if a mobile user is in a moving vehicle, handoff delay could be the user's main concern because of the prompt handoff processing capability in an AP with the lowest load. In the proposed system, the AP's load is a general criterion for AP selection. As indicated in Table 1, the indicators commonly used to reflect the AP's load are contention level, user count, and channel utilization. In [19], contention level was taken as a unique indicator for load balancing, but extra time is required for contention-level probing. User count could reflect the potential AP's load but it is time-consumed for collection. In the proposed system, both the user count and channel utilization are adopted to determine candidate APs' load.

In order to efficiently obtain both figures, a new FHO parameter set is defined in beacon frames, as shown in Figure 6. The new element ID is 99, and the length of information fields is 7 bytes, consisting of the 4-byte IP address of the target DHCP, 2-byte user count, and the 1 byte channel utilization. In addition, a mobile node can easily grasp the current channel utilization and user

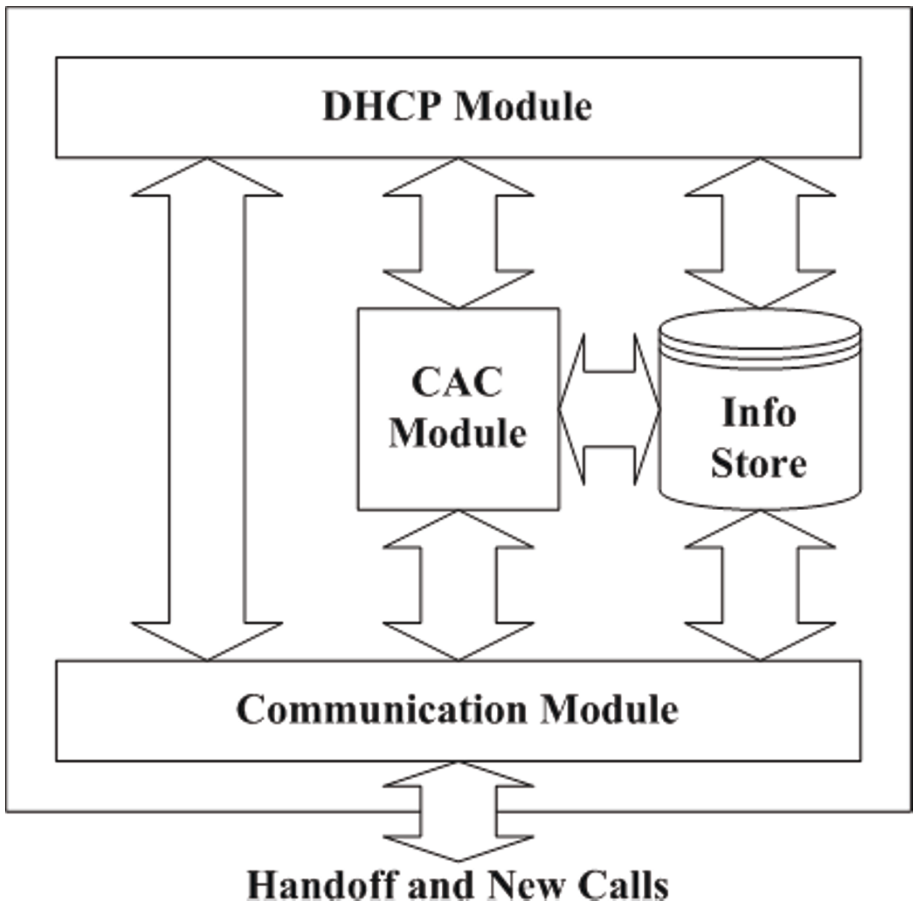

Figure 5 AP architecture.

count of nearby APs through the broadcasted beacons. Furthermore, SNR is also able to be sensed easily. In the proposed AP selection, a MN chooses its target AP by considering the following measures in order: channel utilization, user count, and then SNR. That is to say that the MN will associate with the AP with higher SNR when the first two indicators are the same among candidate APs. Similarly, the AP with lower user count is chosen in the case of same channel utilization.

The benefit of adopting the channel utilization and user count is its low processing cost in determining a target AP. In addition, the AP with the fewer associated users is capable of handling subsequent handoff phases. In our paper, user count-based AP selection is activated only when an AP's SNR is above 30%, that is, the number of users associated with an AP is a primary selection criterion when channel quality is above the threshold. If the signal strength is below 30%, some AP is still selected using the pure SNR-based strategy; however, the AP selection phase would fail under our user count-based mechanism.

3.3. IP address pre-fetch

In the traditional handoff process, IP address configuration setup consumes the most time. In order to achieve FHO, the new IP address of the MN should be determined as quickly as possible. If the new IP address can be obtained before disconnecting with the previous AP, handoff delay will be significantly reduced. Several IP address pre-fetch mechanisms using traditional BOOTP and DHCP relay processes have been proposed [3,10]. In the conventional DHCP relay process, shown in

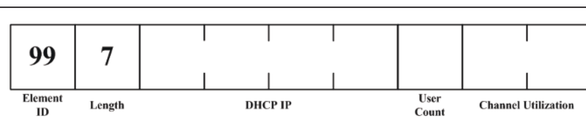

Figure 6 Contents of FHO parameter set.

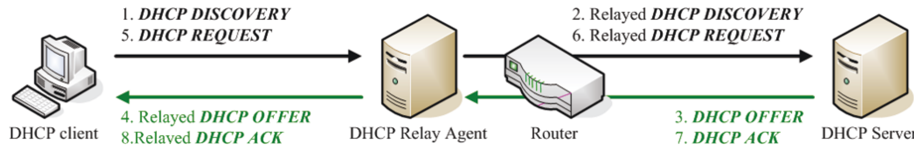

Figure 7 Messages exchanged of traditional DHCP relay.

Figure 7, four DHCP messages, *DISCOVERY*, *OFFER*, *REQUEST*, and *ACK*, are required for exchanging allocated IP address information. The total transmission time is about two RTTs between the MN and the DHCP server. Nevertheless, the time taken by the IP address pre-fetch phase must be short enough to guarantee that an IP address is successfully obtained before disconnect. In this paper, a simplified DHCP relay-based IP address pre-fetch scheme is proposed. In the simplified scheme, due to the target AP having already been discovered, the IP address of the target AP is encapsulated into the *DHCP DISCOVERY* message to inform the local DHCP Module. Once the mobile node receives the *OFFER* message, the MN accepts the assigned IP address and lease unconditionally. The *DHCP REQUEST* and *ACK* messages are simply omitted. Consequently, the time spent for the simplified DHCP relay process, as shown in Figure 8, is reduced to one RTT.

In the proposed FHO scheme, the AP acts as both DHCP server and DHCP relay agent. While the MN is moving away from AP_1 to AP_2, as shown in Figure 9, the MN will receive the beacon from AP_2 embedded within its IP address. Thus, the *DHCP DISCOVERY* message could be relayed via the DHCP module in AP_1. The MN will retrieve an IP address via a *DHCP OFFER* message, once the requirement is admitted by the CAC module. Therefore, the MN could adjust the IP address immediately after its disassociation from AP_1. Since the admission has been granted, all ongoing services could be resumed following the association with AP_2.

3.4. ELFGCP algorithm

The LFGCP is simple and frequently used for CAC. LFGCP favors handoff calls in making channel reservations, which is inefficient when there are more new calls than handoff calls. In this paper, an efficient approach to better use the bandwidth, called ELFGCP, is proposed. In the modified algorithm, both dropping probability threshold (DPT) and blocking probability threshold (BPT) are included to enhance channel utilization. Basically, the new calls will be admitted if the DP is lower than DPT. By including BPT and DPT in the algorithm, channel utilization increases, as described in the next section. The ELFGCP algorithm is as follows:

```
1 if (a handoff call arrives)
2   if (Num_Of_Occupied_Channels <C)
3      accept the call
4   else
5      reject the call
6 if (a new call arrives)
7   if (Num_Of_Occupied_Channels <T)
8      accept the call
9   else if (Num_Of_Occupied_Channels <C) && (DP < DPT)
10      accept the call
11   else if (Num_Of_Occupied_Channels < C) && (BP > BPT) && (random (0,1) <β)
12      accept the call
13   else
14      reject the call
```

The first part of the algorithm, from lines 1 to 5, deals with handoff calls. If there are channels left, that is, if the number of occupied channels is less than the capacity (C), handoff calls are always accepted. The second part, from lines 6 to 14, deals with new calls. If the sum of occupied channels is below the threshold (T), the new call will be accepted; otherwise it is subject to measurement of DP and BP. In lines 9 and 10, if *DP* of handoff calls is low, new calls are accepted because considerable resources are reserved for handoff calls. In lines 11 and 12, the new calls may be approved in a random fashion because BP is exceeded, even though DP is higher than DPT. The purpose of beta is to control the probability of the acceptance of new incoming calls. Since the ELFGCP is a user count-based CAC

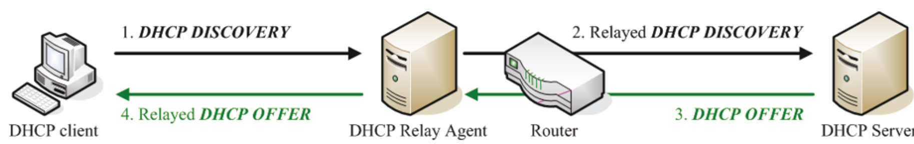

Figure 8 Messages exchanged of simplified DHCP relay.

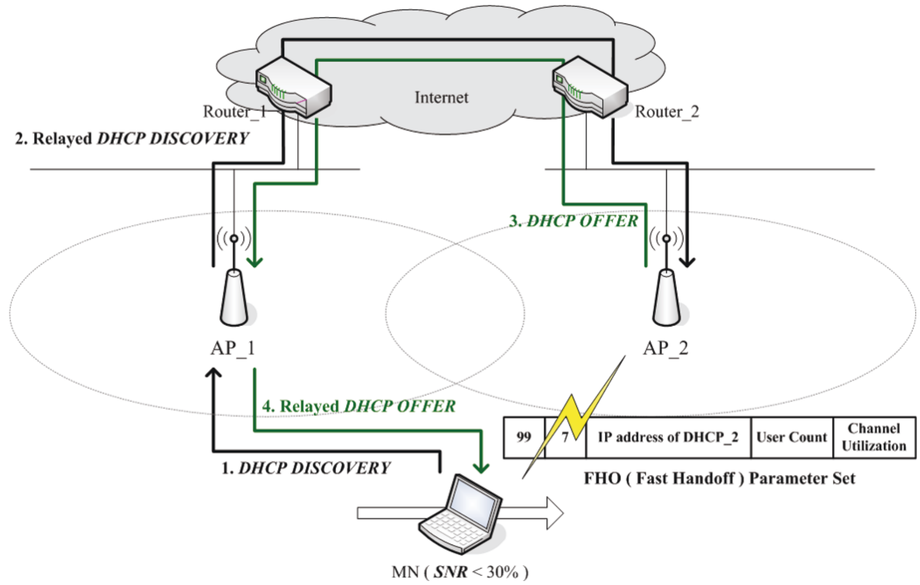

Figure 9 Work flow of proposed IP address pre-fetch.

mechanism, beta should be adjusted dynamically based on user count. Thus, the value of β is set to 1/(users count). With increase in the number of associated users, the difficulty for new calls to get approved also increases. The larger number of associated users reflects the higher acceptance priority for handoff calls in contrast to new calls. Finally, new calls will be rejected if the capacity has been consumed or DP and BP are already balanced.

In most recent CAC research, such as reported in [16], handoff calls are distinguished from new calls. In the proposed design, the CAC Module receives handoff calls via the DHCP module and new calls via the communication module. The detailed operation flow will be described next.

3.5. Operation flow of FHO

Once the target AP has been chosen, both IP address pre-fetch and CAC need to be carried out before the MN disconnects from the serving AP. The overall operation flows of both functions are shown in Figure 10.

When the SNR is below the predefined threshold of 30%, the handoff process is triggered. The MN chooses a target AP from candidate APs, according to their user counts. Since the IP address of the target AP is encapsulated into beacon frames, the MN sends a *DHCP DISCOVERY* with this IP address to the local DHCP module (1). The *DHCP DISCOVERY* is then relayed to the target AP (2). Once the DHCP Module of the target AP receives this message, it acquires the association permission for the MN (3.1). If the CAC test passes (3.2), the DHCP module allocates an IP address for the MN and records the assigned IP address as well as the MN's MAC address in the information store (3.3). The *DHCP OFFER* with assigned IP address and corresponding lease are sent back and eventually forwarded to the MN via the serving AP (4). Only when the MN successfully retrieves the new IP address via the relayed *DHCP OFFER* (5), and the *SNR* is still below the threshold, does it send a *DISASSOCIATION* to disassociate from the serving AP and establish a new association with the target AP.

When the communication module in the target AP receives an *ASSOCIATION REQUEST* (1), as shown in Figure 11, it checks whether the record of the MAC address and assigned IP address of the requested MN exists in the information store (2). If the record is found, the *ASSOCIATION REQUEST* is recognized as a handoff call; otherwise, a new call. When a new call arrives, the *ASSOCIATION RESPONSE* with acceptance notice is sent back (4) if the CAC test is passed, according to the proposed ELFGCP algorithm (3). If the *ASSOCIATION REQUEST* is recognized as a handoff call, the *ASSOCIATION RESPONSE* is immediately sent back (4). Thus, the complete handoff process is accomplished whenever the AP switch phase (T1) and network configuration phase (T4) are completed. Consequently, the handoff delay will be reduced to MAX (T1, T4).

Nevertheless, special care is necessary if association with the serving AP is broken before the MN has successfully retrieved a new IP address. If the *DHCP*

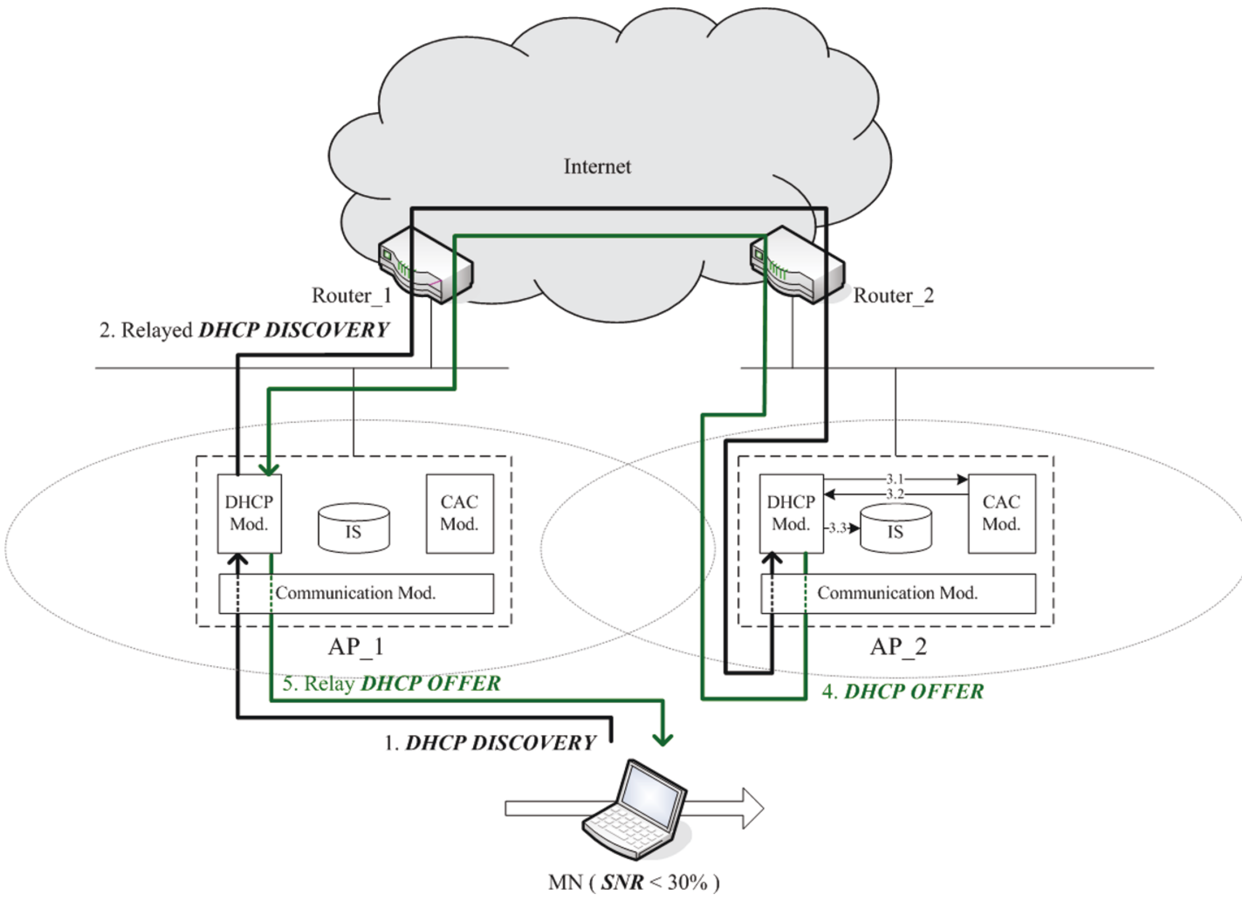

Figure 10 The overall operation flow of IP address pre-fetch and CAC.

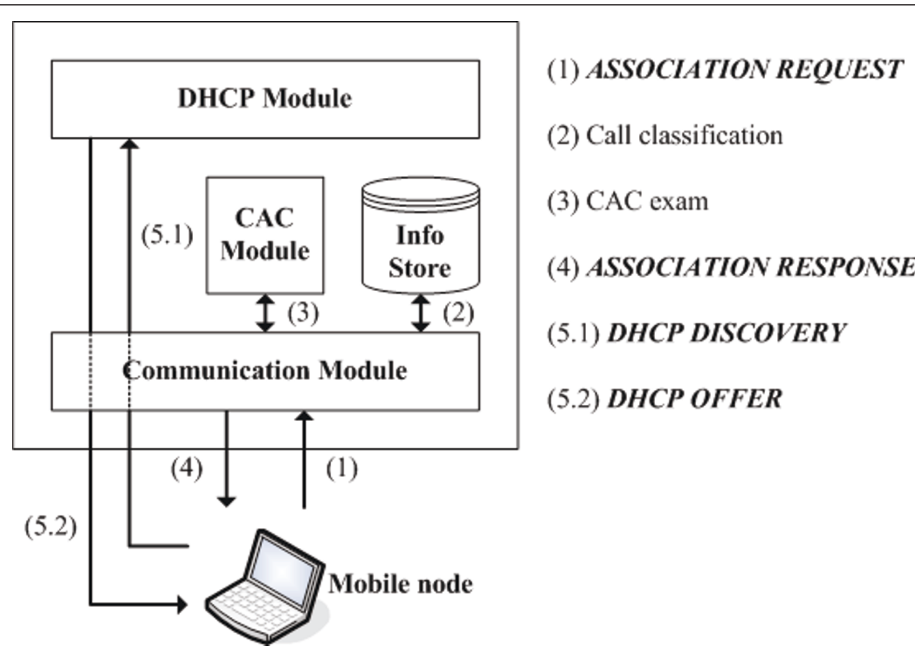

Figure 11 Operation flow of AP switch.

DISCOVERY is sent to the target AP and the admission request is granted, the assigned IP address is recorded in the IS, even though the MN may not be able to receive the *DHCP OFFER*. In such case, the *ASSOCIATION REQUEST* (1) of the MN is recognized as a handoff call because the record can be found in the IS. The communication module will immediately send an *ASSOCIATION RESPONSE* back to the MN (4). The MN will eventually retrieve the previously assigned IP address via the simplified DHCP process (5.1 and 5.2).

The complete operation flow of the proposed FHO process is given in Figure 12. In this handoff process, the disconnect duration is determined solely by the processing of the *ASSOCIATION REQUEST* and sending of the *ASSOCIATION RESPONSE* (red lines). This is because AP selection, CAC, and IP address pre-fetch are finished before disconnect with the serving AP. While the target AP processes the *ASSOCIATION REQUEST* and *ASSOCIATION RESPONSE* (T1), the MN can reset the network parameters (T4). Consequently, the handoff delay will be the optimum value: MAX(T1, T4). With minimized handoff delay, more time is available for processing upper-layer communication protocols, such as delivering the SIP re-invitation to resume the service. How to reduce the minimum delay is hardware dependent and beyond the scope of this paper.

4. Simulation results

Three experiments have been developed using OMNet++ and C++ to investigate the performance of the proposed handoff process. In Section 4.1, the simulation evaluates the throughput comparisons between the

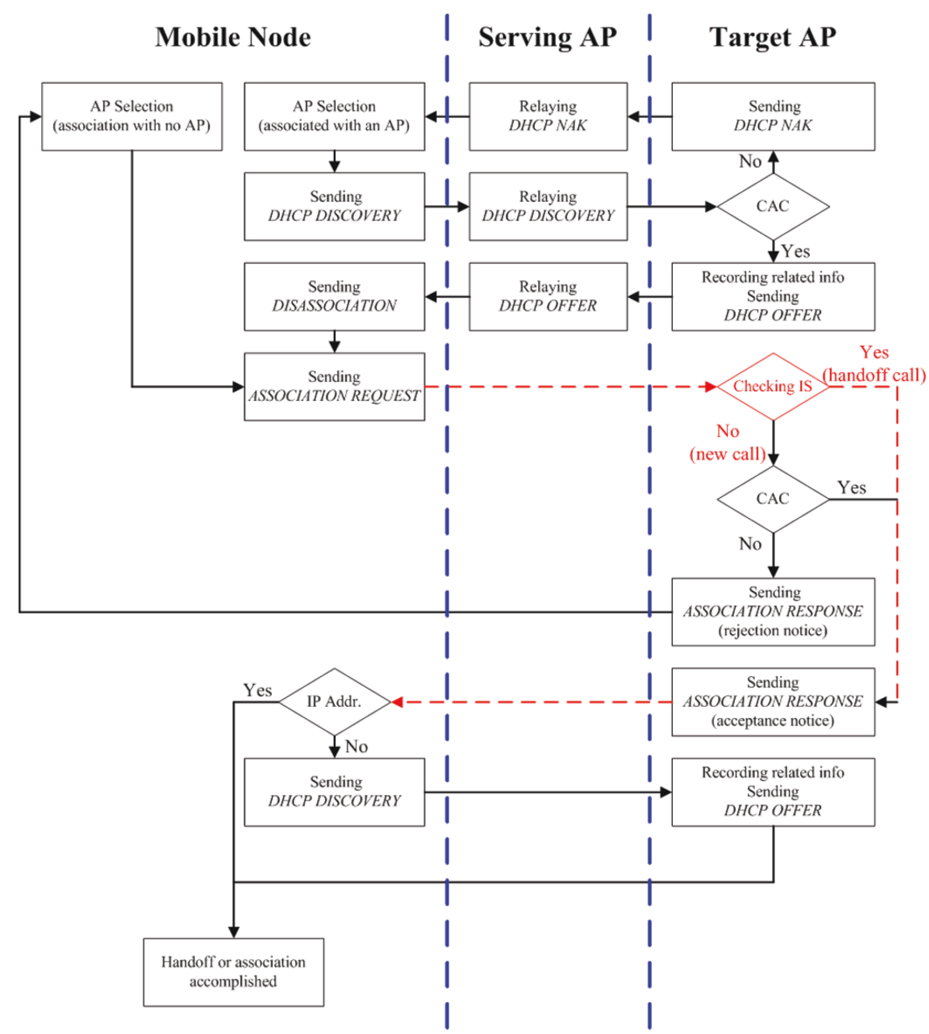

Figure 12 Operation flow of the proposed fast handoff process.

proposed AP selection scheme and the conventional SNR-based approach. The ELFGCP algorithm is emulated in 4.2 to demonstrate the reduction in dropping and blocking probabilities. The overall handoff delay of the proposed system is also reported. The results show that handoff delay can be effectively reduced to MAX (T1, T4), that is, the larger time of either AP switch delay or network configuration latency.

4.1. Throughput measurement under different AP selections

In the proposed AP selection phase, both channel utilization and user-count are adopted for determining the target AP. The commonly used indicator, MPD, is not adopted because of its extra probing duration. Another determination factor of choosing both indicators is because of the easy AP permission in the CAC phase, so that both CAC and IP address pre-fetch can be successfully accomplished before disconnect.

The throughput measurement of adopting either SNR-based or the proposed load-based AP selection was simulated in an area of 500 m^2 that initially contains 63 APs and 2500 MNs. The simulation parameters are given according to the WIFLY service [29] in Taipei City, Taiwan, as listed in Table 2. In the hot spot area, average of 250 APs and 10,000 users exists within 1 km^2. All MNs moved randomly according to the module *Random WPMobility*, as for random waypoint mobility, in OMNet++. Both SNR-based and load-based AP selection strategies are repeatedly simulated. The comparative throughputs for ten simulations are displayed in Figure 13.

The maximum capacity of each AP is set to 255 connections. An *ASSOCIATION REQUEST* from a randomly moving MN is rejected if the AP has already served 255 MNs. As shown in Figure 13, the throughput of load-based AP selection is consistently better than that of SNR-based AP selection in each simulation run. The results reflect that by adopting load-based strategy, the *ASSOCIATION REQUEST* of an MN is mostly accepted to have better overall throughput. Specifically, the average throughput is enhanced by 56%, from 37.2843 to 58.1006 mega bits, with 95% confidence interval ranges from 35.3522 to 39.2164 mega bits and from 49.9551 to 66.2461 mega bits, respectively. The worse throughput of using SNR-based approach is mainly due to the fact that numerous MNs could simultaneously associate with the same AP. When this happens, two undesired situations are possible: (1) an MN may attempt to associate with an already overloaded AP, or (2) there exist a large volume of packet collisions. Either case may result in data re-transmission. Consequently, throughput is reduced. On the other hand, the expanded 95% confidence interval of overall throughput of the proposed scheme reflects that the unstable transmission, which is resulted in the MNs associates with the AP according to the channel utilization and user count instead of the SNR. Thus, throughput is reduced by the increased transmission error.

Table 2 Parameters for throughput simulation

Simulation duration	600 s
Simulation area	500 × 500 m^2
Maximum served users (each AP)	255
Signal coverage	150 m
Average move speed of MNs	10 m/s
Number of APs	63
Number of MNs to be measured	2500
Data transmission rate	100 Kbps

4.2. Performance effects of ELFGCP

To determine the call admission in the proposed CAC algorithm, two auxiliary thresholds, DPT and BPT, are defined. As the smaller probability of failure, either dropping or blocking indicates that more users can access the services, we investigate the failure probability of the proposed CAC algorithm under various new calls to handoff calls ratios in the following simulation. In each experiment, average value of the summed probability is calculated from 12 simulation runs, with the largest and smallest values being excluded. The simulation parameters are given in Table 3.

According to the report from Ramjee et al. [28], the ideal DP should be lower than 0.01. In the following, 0.01 is taken as the threshold for DP. The dropping and blocking probabilities are co-related. That is to say, that only when the DP is acceptable, resources are allocated to new calls. For instance, when the BP is set to be 0.2, if the user count is < 255 and the DP is < 0.01, one out of five reserved calls is made available to a new coming user.

The ratios of new calls to handoff calls in Figure 14a, b, c are set to 4:1, 1:1, and 1:4, respectively. In these figures, the x-axis represents the association time between an MN and an AP, ranging from 60 to 600 s, and the y-axis represents the failure probability. From the simulation results, we found when there are more of new calls, ELFGCP performs much better in the failure probability than what FLGCP does. The improvement is due to that the resource can be flexibly allocated to new calls in the enhanced algorithm, especially when the DP is small. In particular, when the association time is 240 s, the failure probability can be reduced up to 45%, as shown in

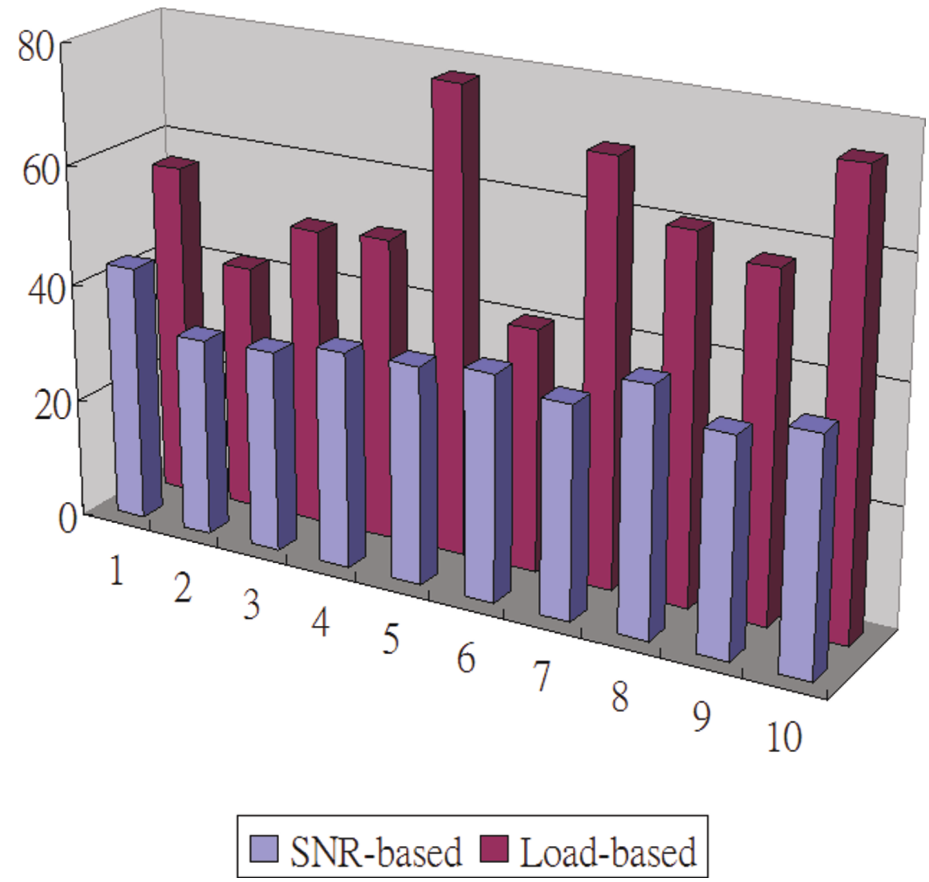

Figure 13 Overall throughput comparisons.

Figure 14a. On the other hand, in the case of either equal call requests or more handoff calls as shown in Figure 14b, c, if the association time is < 240 s, ELFGCP also outperforms FLGCP. When the association time is more than 240 s, both algorithms perform almost identically in the failure probability. Generally, the proposed ELFGCP algorithm provides a better resource utilization than the LFGCP algorithm does, hence the smaller failure probability can be concluded.

4.3. Overall handoff process

To investigate the overall handoff cost, we built an experimental environment consisting of six APs in three groups with a random number of users associating with each AP, as shown in Figure 15. As the designated MN moves, two handoff occurrences are necessary. During the movement, the MN chooses one of two candidate APs from each group. Over a total of 109.3369 s simulation, the connection time, the disconnection time, and the time spent in handoff were measured. The procedure was repeated 22 times and the average times were calculated at each stage, excluding the largest and smallest ones.

Table 3 Simulation parameters

Simulation interval (s)	1000
Capacity	255
Threshold	230
DPT	0.01
BPT	0.2

The total disconnect time includes the time of the MN making the initial association with an AP in Group 1 and the time during the switches (from Groups 1 to 2 and from Groups 2 to 3). Where the MN enters into Group 1, no AP is initially associated with, hence both traditional handoff and the proposed scheme take about 2.9 s to establish the connection, as shown in Table 4. The time spent in each stage as the MN moves in either the traditional handoff or the proposed scheme is presented in Table 5. Since the MN under the proposed handoff scheme attempts to associate with the next AP when the SNR is below 30%, an MN stays in a Group has less associating time. Conventionally, an MN keeps associating with the AP until no radio signal is received. However, the MN in Group 3 under the proposed

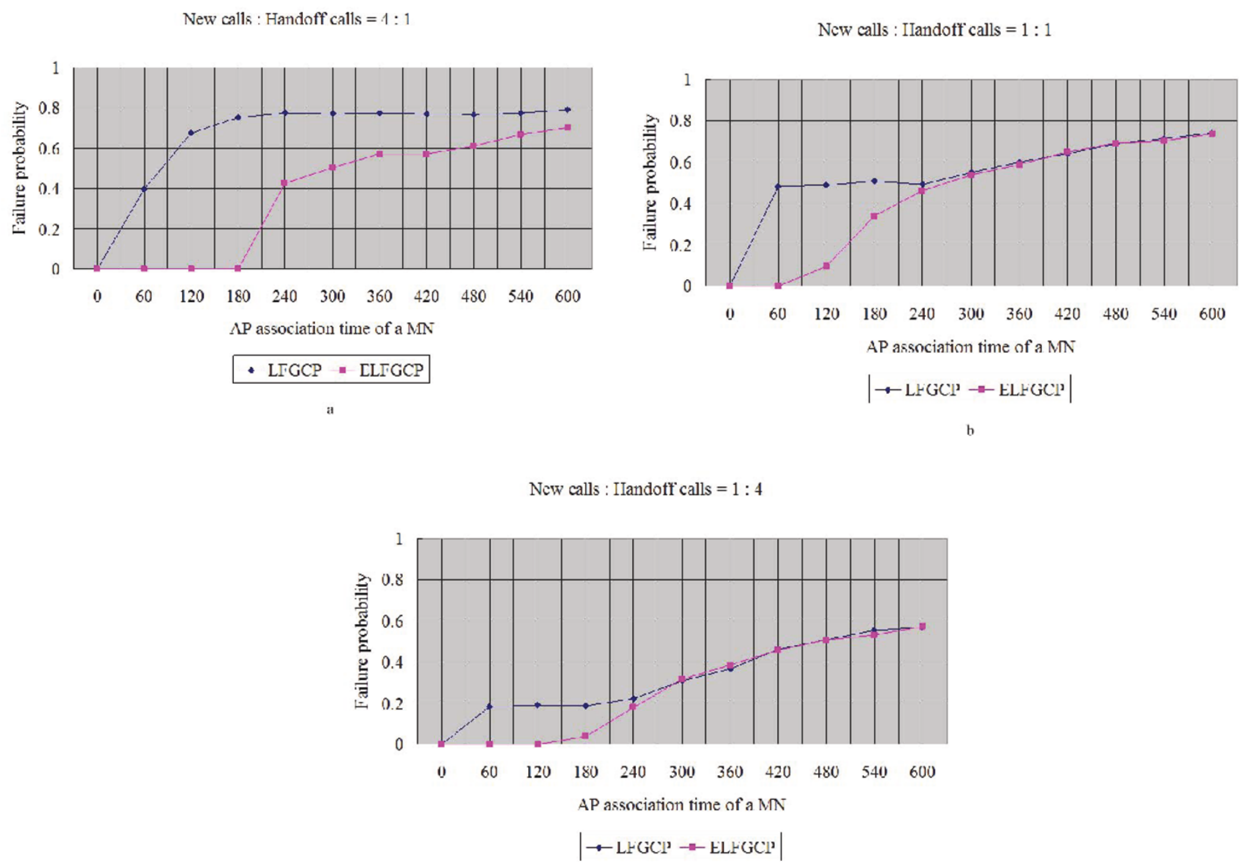

Figure 14 Comparisons between LFGCP and ELFGCP.

handoff scheme cannot discover any available AP even where the SNR is below 30%. Thus, the time that the MN stays in Group 3 is longer than the time it stays in Groups 1 and 2. With the proposed handoff procedure, AP selection, IP address allocation, and CAC are pre-processed prior to disconnect. Accordingly, the handoff durations for both disconnect events, from G1 to G2 and from G2 to G3, are effectively reduced to about 0.004 s, in contrast to more than 2.9 s using traditional handoff.

5. Conclusions

To speed up the handoff process in wireless networks, we present a re-organized scheme of AP selection phase, CAC phase, and IP address re-allocation phase. By pre-processing all possible handoff required operations, the handoff disconnection time can be reduced, while the maximum value of either the AP switch delay or the network configuration time. In the integrated system, the targeted AP is selected by taking into consideration of both the lowest channel utilization and smaller number of associated users. The user count indication for AP selection benefits from the better throughput and lower probability of the call blocking and dropping. And, both call admission and IP address pre-fetch are performed through the simplified DHCP procedure.

In this paper, the ELFGCP algorithm is proposed for CAC, and two additional thresholds, BPT and DPT, are defined for the purpose of evaluating resource utilization. New calls and handoff calls are correlatively treated in the proposed handoff system. The simulation results show that ELFGCP is superior to LFGCP in most situations. Only when there are considerable handoff calls and longer AP association time, both the proposed ELFGCP and LFGCP are almost identical in failure probability performance. The figures indicate that the pre-processing of AP selection, CAC, and IP address pre-fetch can significantly reduce handoff time. Specifically, handoff duration can be reduced to about 0.004 s, as opposed to more than 2.9 s using conventional handoff. In the future, the study of mobile node availability under high-speed motion deserves further exploration. Furthermore, integration of authentication and authorization may also be investigated.

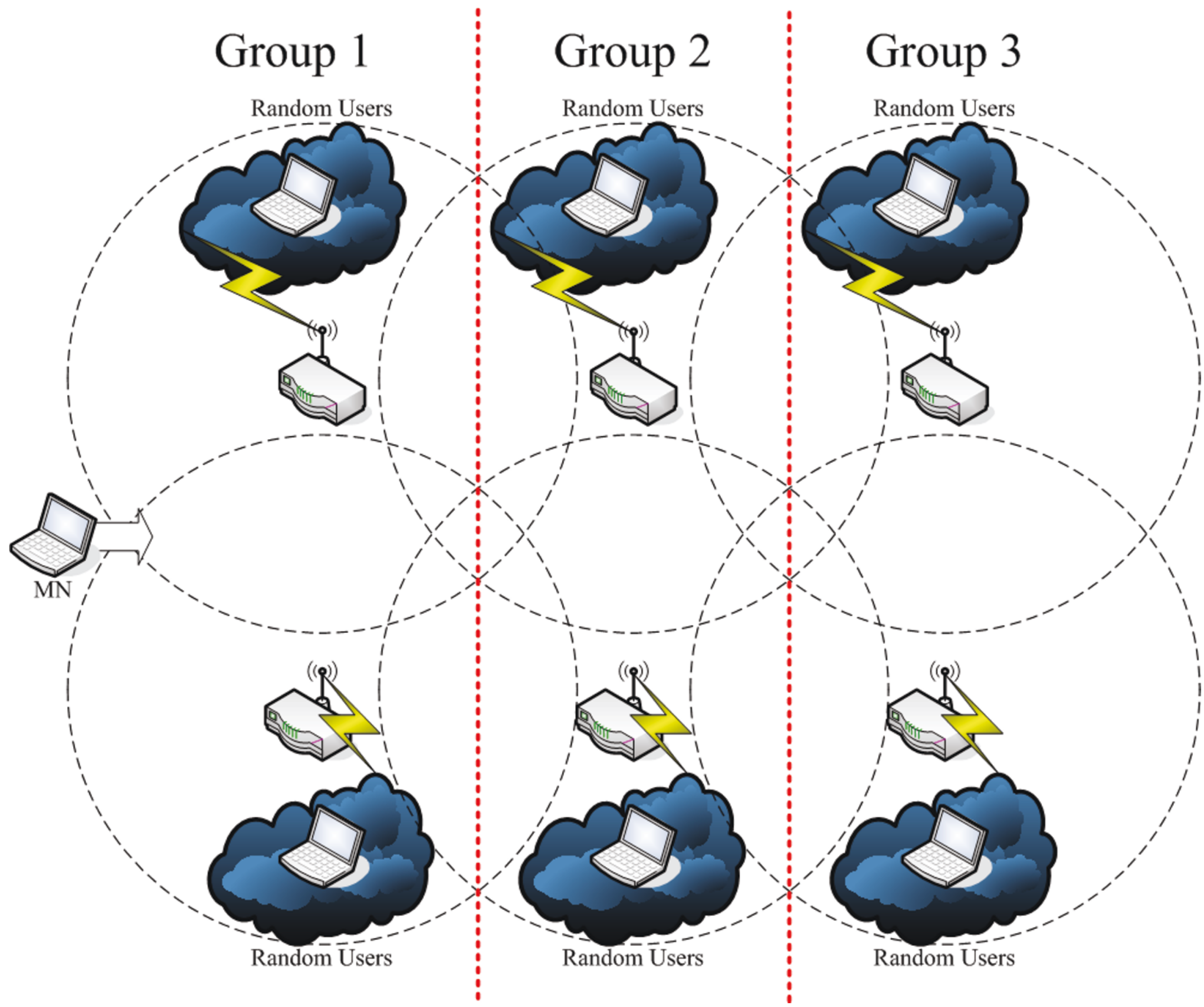

Figure 15 Evaluation environment.

Table 4 Time marks along the MN moves

	Associate with G1	Disassociate from G1	Associate with G2	Disassociate from G2	Associate with G3	Disassociate from G3
Traditional scheme	2.9042	36.3649	39.2766	72.8529	75.7645	109.3369
Proposed scheme	2.9042	31.8786	31.8827	68.3666	68.3708	109.3369

Table 5 Time spent in each stage

	Associate with G1	Stay in G1	Handoff to G2	Stay in G2	Handoff to G3	Stay in G3	Connection time
Traditional scheme	2.9042	33.4607	2.9117	33.5763	2.9116	33.5724	100.6094
Proposed scheme	2.9042	28.9744	0.0041	36.4839	0.0042	40.9661	106.4244

List of abbreviations used

AP: access point; BP: blocking probability; BPT: blocking probability threshold; CAC: call admission control; DP: dropping probability; DPT: dropping probability threshold; ELFGCP: efficient limited fractional guard channel policy; FGCP: fractional guard channel policy; FHO: fast handoff; GCP: guard channel policy; LFGCP: limited fractional guard channel policy; MN: mobile node; MPD: mean probe delay; PHTE: potential hidden terminal effect; RTTs: round-trip times; SIP: session initiation protocol; SNR: signal-to-noise ratio.

Competing interests

The authors declare that they have no competing interests.

Received: 1 February 2011 Accepted: 6 July 2011 Published: 6 July 2011

References

1. IEEE Standard for information technology–telecommunications and information exchange between systems–local and metropolitan area networks–specific requirements–Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999),* June 2007.
2. Dutta A, Madhani S, Chen W, Altintas O, Schulzrinne H: Fast-handoff schemes for application layer mobility management. In *15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications. Volume 3.* ; September 2004:1527-1532.
3. Kim W, Kim M, Lee K, Yu C, Lee B: Link layer assisted mobility support using SIP for real-time multimedia communications. *Proceedings of the Second International Workshop on Mobility Management and Wireless Access Protocols* ; 2004127-129.
4. Yeea YC, Choongb KN, LY Lowa A , Tan SW: SIP-based proactive and adaptive mobility management framework for heterogeneous networks. *J Netw Comput Appl* 2008, **31**:771-792.
5. Rosenberg J, Schulzrinne H, Camarillo G, Johnston A, Peterson J, Sparks R, Handley M, Schooler E: SIP: session initiation protocol. *RFC* June 2002, 3261.
6. Nicholson A, Chawathe Y, Chen M, Noble B, Wetherall D: Improved access point selection. *ACM Mobisys* June 2006.
7. Patrick M: DHCP relay agent information option. *RFC* January 2001, 3046.
8. Wimer W: Clarifications and extensions for the bootstrap protocol. *RFC* October 1993, 1542.
9. Croft WJ, Gilmore J: Bootstrap protocol. *RFC* September 1985, 951.
10. Cha E, Lee K, Kim M: Cross layer fast handoff for SIP. *21st International Conference on Advanced Information Networking and Applications* ; 2007443-450.
11. Fukuda Y, Honjo M, Oie Y: in Development of access point selection architecture with avoiding interference for WLANs. *IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications* ; September 20061-5.
12. Balachandran A, Bahl P, Voelker G: Hot-spot congestion relief and service guarantees in public-area wireless networks. *SIGCOMM Comput Commun Rev* 2002, **32(1)**:59-59.
13. Judd G, Steenkiste P: Fixing 801.11 access point selection. *Proceedings of ACM Mobicom* August 2002.
14. Bejerano Y, Han S, Li L: Fairness and load balancing in wireless LANs using association control. *Proceedings of ACM Mobicom* October 2004.
15. Du L, Bai Y, Chen L: Access point selection strategy for large-scale wireless local area networks. *IEEE WCNC* ; 20072161-2166.
16. Huang KS, Hsieh IP, Kao SJ: Incorporating AP selection and call admission control for seamless handoff procedure. *ICCCE* ; 2008.
17. Vasudevan S, Papagiannaki K, Diot C, Kurose J, Towsley D: Facilitating access point selection in IEEE 802.11 wireless networks. *ACM Sigcomm IMC, Berkeley* 2005.
18. Fukuda Y, Fujiwara A, Tsuru M, Oie Y: Analysis of access point selection strategy in wireless LAN. *Vehicular Technology Conference* 2005.
19. Chen JC, Chen TC, Zhang T, Berg E: Effective access point selection algorithms for handoffs in IEEE 802.11 wireless LANs. *IEEE Global Telecommunications Conference* ; November 20061-6.
20. Case JD, Fedor M, Schoffstall ML, Davin J: Simple network management protocol (SNMP). *RFC* 1990, 1157.
21. Koutsopoulos I, Tassiulas L: Joint optimal access point selection and channel assignment in wireless networks. *IEEE/ACM Trans Netw* 2007, **15**:521-532.
22. Yen LH, Yeh TT: SNMP-based approach to load distribution in IEEE 802.11 network. In *IEEE 63rd Vehicular Technology Conference. Volume 3.* ; 2006:1196-1200.
23. Misra IS: A Banerjee, A novel load sensitive algorithm for AP selection in 4G networks. *Proc International Conference on Computers and Devices for Communications, CODEC04, Calcutta, India* Jan 1–3 2004.
24. Bianchi G, Fratta L, Oliveri M: Performance evaluation and enhancement of the CSMA/CA MAC protocol for 802.11 wireless LANs. *Proceedings of the PIMRC 1996* Taipei, Taiwan; October 1996, 392-396.
25. IEEE Standard for Local and Metropolitan Area Networks. *IEEE 802.16 Standard* 2004.
26. Raychaudhuri D: 4G network architectures: WLAN hot-spots, infostations and beyond. *International 4G Forum, London, UK* May 2002.
27. Ahmed MH: Call admission control in wireless networks: a comprehensive survey. *IEEE Commun Surv Tut* 2005, **7**:49-68.
28. Ramjee R, Towsley D, Nagarajan R: On optimal call admission control in cellular networks. *Wireless Netw* 1997, **3**:29-41.
29. WIFLY . [http://www.wifly.com.tw/].

doi:10.1186/1687-1499-2011-30
Cite this article as: Hsieh and Kao: Handoff optimization in 802.11 wireless networks. *EURASIP Journal on Wireless Communications and Networking* 2011 **2011**:30.