Noname manuscript No.
(will be inserted by the editor)

FastScan - A handoff scheme for Voice over IEEE 802.11 WLANs

Ilango Purushothaman · Sumit Roy

Received: date / Accepted: date

Abstract IEEE 802.11 Wireless LANs are increasingly being used in enterprise environments for broadband access. Such large scale IEEE 802.11 WLAN deployments implies the need for client mobility support; a mobile station has to be "handed off" from one Access Point to another. Seamless handoff is possible for data traffic, which is not affected much by the handoff delay. However, voice traffic has stringent QoS requirements and cannot tolerate more than $50ms$ net handoff delay. The basic IEEE 802.11 handoff scheme (implemented in Layers 1 & 2) only achieves a handoff delay of $300ms$ at best, leading to disrupted connectivity and call dropping. The delay incurred in scanning for APs across channels contributes to 90% of the total handoff delay. In this paper, the FastScan scheme is proposed which reduces the scanning delay by using a client-based database. The net handoff delay is reduced to as low as $20ms$ for IEEE 802.11b networks. We next suggest "Enhanced FastScan" that uses the direction and relative position of the client with respect to the current AP to satisfy the latency constraint in IEEE 802.11a scenarios, which have significantly higher scanning delays due to the larger number of channels. The proposed schemes do not need any changes in the infrastructure (access points) and require only a single radio and a small cache memory at the client side.

1 Introduction

In recent years, IEEE 802.11 WLANs have been widely deployed in campuses, enterprise, commercial centers and also as municipal wireless networks. While WLANs were predominantly designed for data, with the advent of Voice over IP, they are increasingly carrying voice traffic (and possibly multi-media in near future).

The IEEE 802.11 standard [1] specifies two operating modes: infrastructure and ad-hoc mode. In ad hoc mode, the source station establishes communication with the destination station without the help of any fixed infrastructure, in a distributed manner. In infrastructure mode - the most common form of deployments - if a mobile station (STA) wishes to send or receive data, it needs to associate with an Access Point (AP). . The AP and its associated clients form a Basic Service Set (BSS). A set of BSSs form an Extended Service Set (ESS). IEEE 802.11 usually follows the Distributed Co-ordination function (DCF), which uses CSMA/CA with a random backoff algorithm. All data transfers between mobile stations or a client and server in the backbone (Internet) is facilitated through the AP. While this mode has the advantages of minimal configuration and low cost, there are some issues which limit the use of WLANs.

Ilango Purushothaman
Cisco Systems
San Jose, CA 95134
E-mail: ilangop@u.washington.edu

Sumit Roy
Department of Electrical Engineering
University of Washington
Seattle, WA 98195
E-mail: sroy@u.washington.edu

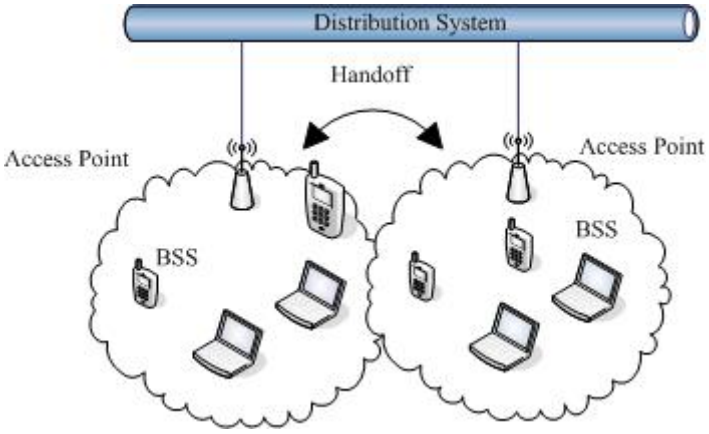

Fig. 1 A typical handoff scenario

Since the access points have a small range (less than 100 m in indoor networks), access points are currently deployed to provide comprehensive coverage first; the optimization of such deployments for performance (aggregate throughput and user delay) still remains work in progress.

Once a station moves out of an AP's range, a 'handoff' is initiated; this is usually triggered by the station's received signal strength falling below *the handoff threshold*. At this point, the station breaks its association with the current AP and starts the handoff process to find a new AP to associate with. Essentially, during the handoff process, the station's communication is suspended and all incoming packets are dropped. Management frames are exchanged between the mobile station, the potential new APs and the old AP before the station is able to associate with a newly selected AP. The gap (delay) in connectivity during the handoff process can last for up to $1second$ (depending on the handoff strategy). For applications with stringent QoS requirements like VoIP, a delay of more than $50ms$ results in a dropped call [2].

A typical handoff scenario in a IEEE 802.11 WLAN is shown in Figure 1.The handoff process is essentially divided into three phases - scanning, re-authentication and re-association. Scanning of channels for APs can be done in two ways - passive and active - according to the IEEE 802.11 standard. Passive scanning which involves the reception of beacon frames by the client in each channel, usually takes up to a second to complete and hence is not favored in time-critical situations like a handoff. Active scanning involves active determination of each channel status by sending probes; this process can be completed within $250ms$ in a typical IEEE 802.11b deployment. Once the best AP is identified, the station authenticates and re-associates with the new AP. The default IEEE 802.11 handoff mechanism involving scanning of all channels provides, at best, a handoff delay of $300ms$ that clearly exceeds requirements for voice traffic continuity.

The proposed FastScan strategy reduces the latency by i) reducing both the number of channels to scan and ii) the number of APs to scan in each channel, by utilizing a client database. The database stores information about the neighboring *best* APs and their corresponding channels, for each entry of the current AP. The client uses the database to select a subset of all the system channels and sends a unicast probe request only to the best AP for each channel in the subset. With the help of this information, broadcast probing is avoided and the client only needs to wait for a single probe response. As a result, the latency is reduced considerably (to as low as $20ms$), thereby avoiding a dropped voice call.

Enhanced FastScan builds on the initial idea, and further uses direction and relative position of the client with respect to the current AP to divide the current BSS area into different sub-sectors. Using this information, the handoff latency can be reduced even further in the case of IEEE 802.11a scenarios and provides more precise handoffs in IEEE 802.11b scenarios.

The rest of the paper is organized as follows. Section II presents the IEEE 802.11 background and an overview of the related work. A detailed explanation of the proposed schemes is provided in Section III. NS-2 IEEE 802.11 implementation and the simulation results are discussed in Section IV, followed by the conclusions in Section V.

2 Background and Related Work

The basic handoff process in IEEE 802.11 consists of three phases - Scanning, Authentication and Re-Association.

First, the mobile station has to determine that it is moving out of range of the AP and initiates handoff. This is usually done through monitoring the signal-to-noise ratio (or received signal strength) and detecting a downward crossing of a pre-defined handoff threshold; the channel scanning phase then starts. Velayos et al. [3] suggest for example, that three consecutive transmission failures is enough to trigger a handoff. Handoff can also be detected through missed beacons, if the station is not sending or receive any data packets.

In passive scanning, the station listens for beacon messages on every channel. Thus the mobile node can create a candidate set of APs prioritized by the received signal strength indicator (RSSI) and select the strongest AP. But passive scanning is usually not favored during a handoff scenario, as the station has to wait for at least one *BeaconInterval* (usually $100ms$) to get most or all of the beacons in that channel. The passive scanning delay T_{scan} is given by

$$T_{scan} = N * BeaconInterval + N * ChannelSwitchTime$$

where *ChannelSwitchingTime* is the time taken to switch the radio from one channel to another and N is the number of channels in the IEEE 802.11 spectrum.

The other parameters that contribute to the scanning delay are described below.

- *Contention Time* : This delay due to CSMA deferrals introduced by the Distributed Co-ordination function in IEEE 802.11 and includes the backoff delay. Since the AP and management frames do not enjoy any kind of priority in IEEE 802.11a/b networks, this adds to the total scanning delay.
- *ChannelSwitchTime* : This delay includes the time to switch to a new frequency, synchronize and start demodulation. The value has been observed to vary from $5\,ms$ (Atheros NIC) to $10\,ms$ depending on the vendor implementation.

In IEEE 802.11b networks with 11 channels, using 5 ms as the value for *ChannelSwitchingTime* yields delay of $1050ms$ with passive scanning. This is clearly not suited for delay-sensitive voice traffic and hence active scanning is the preferred method for scanning. It is initiated by the station which does not have to wait to receive any beacons. The station broadcasts probe requests evoking probe responses from the APs thereby reducing the time spent on each channel during scanning. Probe responses are very similar to beacons frames and carry the same information about BSSID, timestamps, and supported data rates. The handoff process timeline in this case is depicted in Figure 2. The active scanning procedure is started by sending out a probe request. The station starts a probe timer at this instance of time, waiting for a probe response. The parameters that affect active scanning are as follows:

- *MinChannelTime* represents the minimum duration a station must wait after sending out the probe request and before deciding that the channel is empty (devoid of APs). If the station receives a *CCA busy* indication before the probe timer reaches zero, the station assumes that one or more APs are operating on this channel. Accordingly, the timer is reset to a value *MaxChannelTime*. The probability that a *CCA busy* before *MinChannelTime* is not a probe response from an AP, is ignored. If no *CCA busy* response is received within *MinChannelTime*, the node can switch to the next channel, deciding that the channel is devoid of APs. This is a tunable parameter and various studies [4] and [3], have suggested values between $1 - 7ms$. *MinChannelTime* should be short enough to catch the first *CCA busy*, to decide whether a channel is empty or not.
- *MaxChannelTime* represents the maximum amount of time a station has to wait to collect probe responses in case the channel is deemed busy. The value of *MaxChannelTime* should ensure that probe responses from most or all of the APs are collected before moving on to the next channel; it's choice thus depends on the AP density in each channel. Studies [4] and [3] have shown that a *MaxChannelTime* of $11ms$ suffices in practice to capture most or all probe responses. This value should be naturally minimized to reduce the scanning phase, but it may result in candidate APs being missed.

Hence, the station waits for *MinChannelTime* on empty channels (devoid of APs) and for *MaxChannelTime* on busy channels.

The total $ProbeWaitTime$ or T_{probe} for all channels is thus bounded by

$$N * MinChannelTime \leq T_{probe} \leq N * MaxChannelTime$$

Hence the active scanning delay T_{scan} is given by

$$T_{scan} = T_{probe} + N * ChannelSwitchTime$$

With values of $5ms$ and $11ms$ for $MinChannelTime$ and $MaxChannelTime$, the total scanning delay for active scanning is found to vary from $105ms$ to $210ms$. Hence, active scanning is much faster than passive scanning and is more widely using in practice for handoff situations.

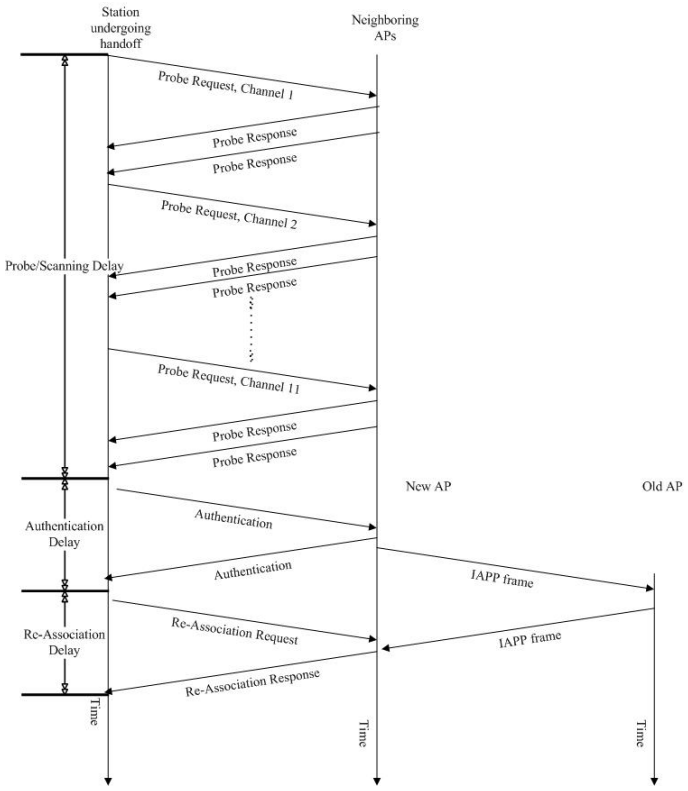

Fig. 2 Basic IEEE 802.11 handoff timeline

Once scanning is over and the best AP has been identified, the station is ready to start the authentication process. Authentication involves the exchange of authentication frames to determine the station's credentials to gain access in the new BSS. With open system authentication, the station sends only an authentication frame, and the access point responds with an authentication frame response indicating acceptance (or rejection). The shared key authentication has more security as it involves encrypting and decrypting a challenge text using a key like WEP. The authentication delay T_{auth} is usually around $5-8ms$, depending on CSMA deferrals.

To send and receive data in the new BSS, the station has to complete re-association with the new AP, after authentication. Re-association enables the AP to allocate resources for and synchronize with the station. Re-association begins with a re-association request to the new AP, containing the old AP's MAC address as the BSSID. The new AP responds with a re-association response, after agreeing on data rates compatibility. Since the new AP knows the old AP's MAC address, it requests state information about the mobile station through the Distribution System. This state information about the station includes

authentication credentials and Inter-Access Point Protocol (IAPP) is usually used for Access Point communication [5]. The re-association delay T_{assoc} is usually $5-8ms$ excluding the IAPP delay T_{IAPP}, which is around $5-10ms$.

$$HandoffDelay = T_{scan} + T_{auth}s + T_{assoc} + T_{IAPP}$$

Hence, the aggregate handoff delay equals $200ms$ on average, in an IEEE 802.11b deployment using 11 channels, when active scanning is considered. If passive scanning is used, the handoff delay is well over a second. As can be seen, the scanning delay takes up almost 90% of the handoff latency. Clearly, this is not good enough for voice applications.

Many handoff techniques have been suggested to reduce the handoff delay. Most of the handoff techniques in literature focus on reducing the scanning delay (active or passive). With respect to active scanning, reducing the scanning delay is achieved either by reducing the number of channels to be scanned or by reducing the probe wait time in each channel, by reducing the number of APs.

Mishra et al. empirically studied the handoff process in [4] for indoor WLAN networks and analyzed the various factors that contribute to the handoff latency. Velayos et al. in [3] also studied the handoff process and suggested techniques to reduce the handoff latency time. In the SyncScan approach [6], the station keeps track of the best AP in its neighborhood by scanning channels periodically. By synchronizing the APs and forcing the APs to transmit the beacon signals in a synchronized manner, the station need not wait for the full beacon interval. Hence the handoff latency is essentially reduced to a few milliseconds, consisting of only the authentication and re-association delays. However, the cost involved with this approach is a regular suspension of communication equal to at least the double the channel switching delay which depending on the hardware, which can exceed 10 ms occurring every SyncScan period. The other problem is synchronization of APs which involves considerable hardware re-configuration.

In the Neighbor graphs approach [7], extra functionality is implemented at both clients and APs that is used to infer the WLAN topology and reduce total handoff overhead to 30-40 ms. This is based on the neighbor reports generated by the APs - however, this has no area-specific (direction based) neighbor AP information, which would be useful in IEEE 802.11a deployments, where there are a lot of neighboring APs operating on non-overlapping channels. This approach, as the previous, involves changes at the APs, which would incur significant cost given the large number of legacy WLANs.

In [8], the total *ProbeWaitTime* was avoided in the scanning phase after sending out the probe requests in all the channels, by receiving all the probe responses from the candidate APs via the old AP. This reduces the handoff delay to an average of 100 ms, which is not suitable for voice traffic. Hence, another scheme was presented to satisfy the voice constraints, by adaptively distributing the total handoff latency. Based on average change in RSS, the station predicts the amount of time it takes before the handoff threshold is reached. During this time period, the normal data transfer is adaptively interleaved with scanning. This scheme reduced the handoff delay to an average of 50 ms. However, there are suspensions in normal communication during the period when scanning is intermittent. Also, both schemes scan all the channels in the IEEE 802.11b spectrum, which is not advantageous for reduction of handoff latency.

In the MultiScan approach [9], Mishra et al. proposed a *multi-radio* solution to the hand off problem. By using a secondary radio dedicated to the scanning and res-association process while the primary radio is used for communication, the handoff delay is almost completely eliminated. When a handoff is needed, the secondary radio is ready with the information and just re-associates with the new AP, effectively becoming the new primary radio. Packets are now transmitted by the secondary radio while the primary radio starts scanning. Though this approach eliminates latency using multiple radios, it is not suitable for hand-held 802.11 devices in which power consumption is critical. A second radio continuously scanning channels while the first radio used for communication, leads to significant increase in battery power consumption.

Among industrial vendors, Cisco 7920 Wireless IP phone's Layer 2 roaming mechanism [10] achieves a handoff delay of less than 100 ms. Cisco 7920 phone initiates a re-association process with a different access point if the phone does not receive three consecutive beacons from the existing access. The 7920 also periodically scans for better access points and maintains a list of potential access points. The decision to handoff is also affected by the utilization factor, which is broadcast by each access point in its beacon. As a result, the phone can avoid attempting to re-associate with access points that have high utilization (load in terms of number of clients and/or traffic) in their BSSs and may not be able to effectively support voice traffic.

Meru Networks and Extricom [11] and [12] have taken a different approach: there is no actual roaming. In their architectures, each AP shares the same BSSID (MAC address). Thus the client effectively sees one ‘ super AP while the WLAN infrastructure manages which AP communicates with a given client. The roaming delay is effectively reduced to around 50 ms.

IEEE 802.11r [13] is a proposed amendment to the IEEE 802.11 standard to permit connectivity aboard vehicles in motion, with fast handoffs from one base station to another managed in a seamless manner. The main objective of this task group is to reduce the roaming delay to less than 50 ms, to enable VoIP over WLANs.

In summary, our proposed FastScan scheme uses a client-based database, which stores information about the neighboring APs and channels, to reduce the number of channels and APs to be scanned. This scheme is able to bring down the handoff latency to less than 50 ms for IEEE 802.11b scenarios. The Enhanced FastScan scheme uses direction and relative position information of the client in addition the neighboring AP and channel information, to divide the BSS into different areas, reducing the handoff latency to sub-50 ms even for IEEE 802.11a scenarios. Both schemes require changes only on the client side and does not need any configuration on the existing infrastructure (APs). Also, both approaches are suitable for current single-radio 802.11 clients and only a small cache memory is needed at the client to implement the FastScan and Enhanced FastScan databases.

3 FastScan and Enhanced FastScan

The FastScan scheme reduces the handoff delay by utilizing a client-based database, which stores information about the neighboring best APs and their corresponding channel numbers. The Enhanced FastScan scheme uses direction and relative position of the client with respect to the current AP to sectorize the current BSS into different areas. Using this additional information, the handoff delay can be reduced even further in the case of IEEE 802.11a scenarios to meet voice handoff latency constraints.

3.1 FastScan

In this scheme, a client side database is used to reduce the number of channels to scan. This scheme also forces the station to search for only one AP per neighboring channel, called the ‘best’ AP. This is achieved by storing and using information about candidate APs in each channel in the database. These entries are stored corresponding to the AP that the station is currently associated with. Since neighboring APs operate on orthogonal channels to minimize adjacent channel interference (e.g. only 3 such channels for IEEE 802.11b), each BSS is typically surrounded by a few APs operating on a subset of the total number of channels. Clearly, there is no need to scan all 802.11b channels, as some channels would just be devoid of APs. Based on previous scans, information about the best AP only is stored for each neighboring channel.

Handoff is initiated when there are three successive transmission failures as in [3]. When handoff starts, only the APs in the channels stored in the database corresponding to the current AP are probed and not all channels in the IEEE 802.11 spectrum. Hence the number of channels to scan is reduced considerably. Since the MAC addresses of only the best APs in each channel are stored, the number of APs to scan in each channel is reduced to just one. An *unicast probe request* is sent to the best AP in each channel, thereby reducing the *Probe Wait Time*, i.e, the time spent on each channel. The best AP, based on RSSI measurements of the probe responses received, is selected for authentication and re-association. The FastScan handoff timeline is depicted in Figure 3. This strategy can reduce the handoff latency to as low as $20ms$ for IEEE 802.11b scenarios.

The FastScan database can be initialized by allowing the mobile station to roam across all the BSSs in the ESS. The best AP found in each neighboring channel is stored along with the channel number, corresponding to the AP that the mobile station is currently associated with. The specific steps for initiating the database are detailed below.

– *Basic Active Scanning of Non-Overlapping channels*: The initial roaming is accomplished by making clients roam across the entire ESS and collecting neighboring AP and channel information for each AP’s BSS. The basic IEEE 802.11 active scanning method is used, where only non-overlapping channels are scanned. The database is initially built by making the client roam into each of the BSSs to create an entry for each BSS, thereby triggering handoffs. The active scanning method is used to collect all

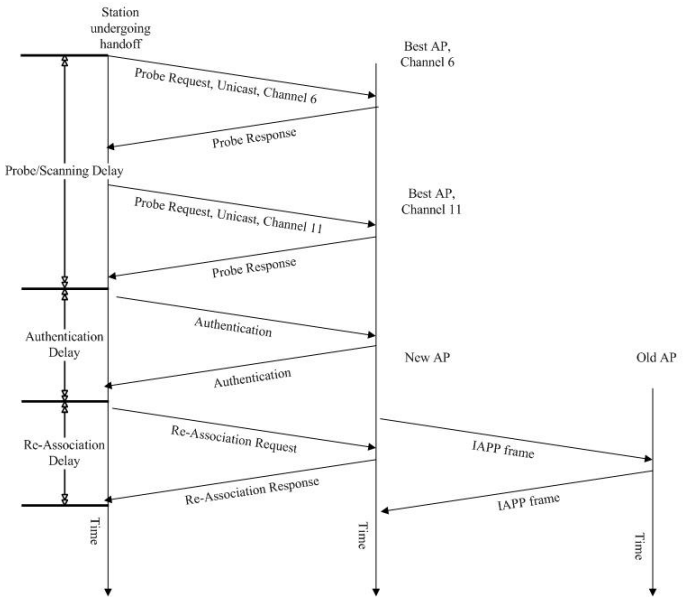

Fig. 3 FastScan handoff timeline

probe responses and build the database based on the strongest APs heard per non-overlapping channel. Since the database would contain only non-overlapping channels, only these channels are scanned (for example 1, 6 and 11 for IEEE 802.11b). Hence, the handoff delays during this initial roaming is much lower than the basic handoff scheme. Even then, voice traffic might suffer during this initial roaming, but this is a one-time overhead only. The time for which the initial roaming is done, depends on the speed of the mobile client and the size of the WLAN deployment used.

- *Beacon Reception* : Beacon frames can also be used to build and constantly update the database entries.
- *Listening to Re-Association Requests* : A mobile station can eavesdrop on Re-Association requests sent by other mobile stations in the same BSS. The Re-Association Requests to the new AP will have the old AP address as one of its fields. If channel information of old AP is also included, mobile stations can use this information to build and update the database quickly.

The database is constantly updated through beacon reception and listening to re-association requests so that when a handoff is needed, the most recent best AP per neighboring channel is probed. By scanning only the neighboring channels and by sending out an unicast Probe Request to the best AP in each neighboring channel, scanning latency is reduced to as low as $15ms$ in an IEEE 802.11b scenario. The authentication and re-association delays are in the order of $5ms$ each. Hence, the total handoff delay is reduced to $20ms$ in most scenarios, lower than $50ms$ constraint imposed by the voice applications. A typical handoff scenario in an IEEE 802.11b environment is shown in Figure 4. The relevant FastScan database for the example scenario is given below in Table 1, with the highlighted entry showing the current BSS.

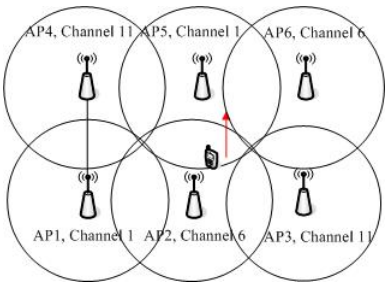

Fig. 4 A typical IEEE 802.11b handoff scenario

Table 1 FastScan database for example IEEE 802.11b scenario

Current AP, Channel	Best APs and channels	
AP 2, Ch 6	*AP 5, Ch 1*	*AP 3, Ch 11*
AP 1, Ch 1	AP 2, Ch 6	AP 4, Ch 11
AP 3, Ch 11	AP 6, Ch 6	AP 5, Ch 1

In a IEEE 802.11b network, there are only 2 neighboring channels as contenders for best AP entries for each current AP in the database. However, storing the best AP only for each neighboring channel comes at a price - the best AP entries are for the entire BSS and are not area-specific. When a mobile station is handing off, there is a probability that these two best AP contenders might not lie in the direction of the mobile client's trajectory, even if the database is constantly updated. In some instances, the best APs might be operating in the direction in which the mobile station is traveling but those entries might not be present in the database. This leads to wrong handoffs, thereby triggering subsequently corrective handoffs within a short time. Clients using FastScan databases usually suffer only a few wrong handoffs in normal WLAN deployments. This is due to the fact that the FastScan database is constantly updated by listening to beacons and re-association requests of other clients. Hence, database entries with old best AP candidates should usually be refreshed with new candidates. However, there is a probability that these database updates do not happen in time and hence a few wrong handoffs are likely to occur. Nonetheless, handoff delays associated with the wrong handoffs and their corrective handoffs will typically occur within the voice delay bound limits.

To counter the wrong handoff problem in FastScan scheme and minimize wrong handoffs, a new *failsafe* mechanism is implemented. When a client is moving from one region to another and the best AP candidates are not in the client's trajectory, the client starts the handoffs and sends out unicast probe requests. The mobile client will most likely hear weak probe responses from these far-off, wrong AP candidates, sometimes hearing only one weak probe response. In rare worst case scenarios, the client does not hear even a single probe response. Instead of starting authentication with a far-off AP and completing a wrong handoff, the client does cross-referencing in its database to determine the next best AP that is close to its current AP and starts authentication with that AP. The mechanism is explained below:

1. During handoff, if the probe responses heard have weak RSSI (controlled by a tunable threshold) or are less than two in number, do not start authentication with the weak AP candidate. These weak APs are termed as failed APs. Start the failsafe mechanism.
2. Search the entire database (for other stored BSSIDs) and short-list those entries which includes the client's current AP, as a neighboring best AP. This is a list of the APs that has picked our own current AP as a best AP. Essentially, we are zooming in on APs neighboring our current AP. This list should have at least one of the failed APs that was tried in Step 1.
3. In this list, find one of the failed AP entries and check the other neighboring best AP candidate (say AP X) that is stored against this failed AP. This other neighboring AP X is most probably the next best AP candidate. There is a very good chance that this AP X is very close to the client's current AP, considering that the failed AP is close to both current AP and AP X, by association. Thus the client does a single intelligent hop, through the weak AP entry. The accuracy of this hit (AP X) is enhanced further, if AP X itself is in the shortlist (meaning that AP X itself has picked the client's current AP as a neighbor).
4. If the failed APs don't turn up in the shortlist, the client picks the first AP from the list, without doing the hop. The probability of such a situation is very rare as failed APs normally comes up in the short list.
5. Start authentication with this selected best AP.

By the above mechanism, a wrong handoff is mostly avoided and the client gets associated to the right AP. Storing "second best APs" per channel will take care of wrong handoffs but will result in significant increase in database size, creating cache constraints on the client. This failsafe mechanism intelligently cross-references only the existing entries in the database and does not add new entries. The utility of this mechanism and the problem of wrong handoffs are best illustrated by the IEEE 802.11b dense deployment shown in Figure 5. If the FastScan database is not updated with the latest information for such a scenario, the failsafe mechanism would kick in and select the best AP, as shown in Table 2. In this case, the client is moving down from AP5 to AP2 but has tried only AP3 and AP6 (failed APs). The client searches the

database and sees that the failed AP3, has AP2 (AP X) and AP5 (current) as the best APs. AP2 is then selected as the AP candidate. For confirmation, the client also checks AP2's best APs (which are AP5 (current) and AP1).

Table 2 FastScan database for example wrong handoff scenario

Current AP, Channel	Best APs and channels	
AP 5, Ch 1	*AP 6, Ch 6*	*AP 3, Ch 11*
AP 6, Ch 6	AP 5, Ch 1	AP 3, Ch 11
AP 3, Ch 11	*AP 2, Ch 6*	AP 5, Ch 1
AP 2, Ch 6	*AP 5, Ch 1*	AP 3, Ch 11

This failsafe mechanism will not be used by a mobile client in most normal WLAN deployments, as there are constant updates to the database. The worst case scenario, where there are no probe responses, is very rare, but has now been addressed by the FastScan scheme. This is simulated and discussed in the next section. Apart from the problem of a few wrong handoffs, FastScan is able to satisfy voice traffic constraints only for IEEE 802.11b scenarios with heavy loads. Increasing the number of non-overlapping channels however, will increase the handoff delay over the $50ms$ threshold, for example in IEEE 802.11a networks where there are 8 non-overlapping channels. To address this problem and to completely eliminate the wrong handoffs, the Enhanced FastScan scheme is proposed.

3.2 Enhanced FastScan

Using Wi-Fi positioning algorithms and knowledge of APs co-ordinates, relative position and direction information can be added to further expand the database and select the APs more precisely, specific to the actual location of the mobile station. By this enhancement, instead of storing just one best AP per neighboring channel, more 'best' APs can be stored depending on the direction in which the mobile station is traveling and the relative position of the mobile station with respect to the current AP.

The relative position of the mobile station with respect to its current AP, can be obtained by using its own approximate co-ordinates (using Wi-Fi positioning) and knowledge of the co-ordinates of the nearby APs can be obtained by broadcasts by the APs, using secure authentication. By this method, each BSS area can be split four ways into North-East, North-West, South-East and South-West of the current AP and a separate entry for the best AP is entered for each area. Hence for each current AP entry, the mobile station has different best APs in each channel, depending on the direction (area) of the BSS in which the mobile station is now located. The best neighboring APs in Channel 6 and 11, for example, might be different for different areas of the current BSS.

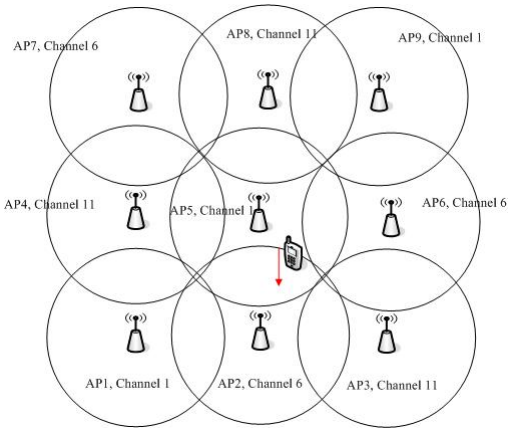

Fig. 5 A dense IEEE 802.11b deployment

Table 3 Enhanced FastScan database for the example dense IEEE 802.11b scenario

Current AP, Channel, Direction(Area)	Best APs and channels	
AP 5, Ch 1, NE	AP 6, Ch 6	AP 8, Ch 11
AP 5, Ch 1, NW	AP 7, Ch 6	AP 4, Ch 11
AP 5, Ch 1, SE	*AP 2, Ch 6*	*AP 3, Ch 11*
AP 5, Ch 1, SW	AP 2, Ch 6	AP 4, Ch 11

A typical IEEE 802.11b dense deployment is shown in Figure 5, and the relevant Enhanced FastScan database is shown in Table 3, where there are different best neighboring APs for each direction (area), thereby providing precise handoffs. In this example, the mobile station is in the South-East area of the current AP and is moving away from it. The station accesses the South-East entry of the current AP and scans those specific neighboring channels and the corresponding best APs, close to the South-East part of the current BSS.

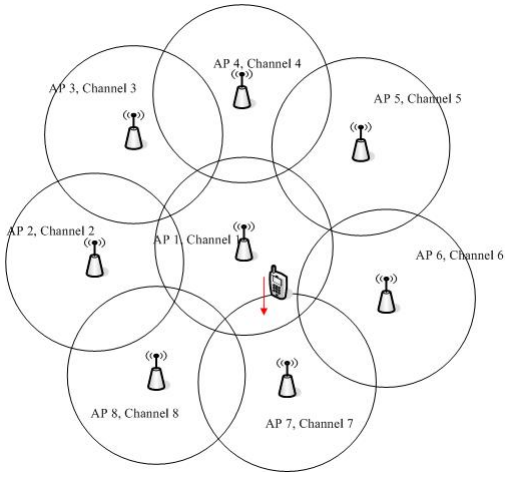

Fig. 6 A typical IEEE 802.11a deployment

Enhanced FastScan scheme is most useful for meeting the voice constraints in IEEE 802.11a scenarios. IEEE 802.11a has 8 non-overlapping channels in its 5 GHz band. Most IEEE 802.11a BSSs are deployed so that neighboring BSSs operate on these non-overlapping channels. A simple FastScan mechanism in an IEEE 802.11a without direction information can result in so many entries per current AP, as the mobile station in a BSS tends hear as many as 7 "best" APs on seven different neighboring non-overlapping channels. However, using an Enhanced FastScan database, a mobile station has to scan only 2 to 3 channels and best APs, as they are specific to that area of the BSS. Also, since the Enhanced FastScan database is area-specific and very precise, wrong handoffs are mostly eliminated.

Table 4 FastScan database for the example IEEE 802.11a scenario

Current AP, Channel,	Best APs and channels		
AP 1, Ch 1	AP 2, Ch 2	AP 3, Ch 3	AP 3, Ch 3
AP 1, Ch 1	AP 5, Ch 5	AP 6, Ch 6	AP 7, Ch 7
AP 1, Ch 1	AP 8, Ch 8		

A typical IEEE 802.11a deployment is shown in Figure 6. For the same IEEE 802.11a deployment, relevant entries of the FastScan and Enhanced FastScan database are shown in Table 4 and Table 5 respectively, with italicized entries showing the current BSS. As can be seen, a mobile station going south from the south-east area of the current BSS, using a FastScan database without direction information, has to scan all the 7 neighboring channels and best APs thereby producing a handoff delay of well over $50ms$.

Table 5 Enhanced FastScan database for the example IEEE 802.11a scenario

Current AP, Channel, Direction(Area)	Best APs and channels		
AP 1, Ch 1, NE	AP 4, Ch 4	AP 5, Ch 5	AP 6, Ch 6
AP 1, Ch 1, NW	AP 2, Ch 2	AP 3, Ch 3	AP 4, Ch 4
AP 1, Ch 1, SE	*AP 5, Ch 5*	*AP 6, Ch 6*	*AP 7, Ch 7*
AP 1, Ch 1, SW	AP 2, Ch 2	AP 7, Ch 7	AP 8, Ch 8

With direction information stored, the mobile station has to scan only 3 neighboring channels and the best APs in those channels. This helps to keep the handoff delay under the $50ms$ VoIP constraint.

4 NS-2 Implementation, Simulations and Analysis

The IEEE 802.11 model implemented in current ns-2.32 distribution is very limited in its feature set. It implemented only the ad-hoc mode of 802.11 and did not include infrastructure mode support. Access Point capabilities, Passive and Active Scanning, Authentication, Association, handoff support were not implemented. Also, ns-2.32 could support only a single channel network. Hence, a full-featured IEEE 802.11 infrastructure mode model was implemented to support WLAN simulations in ns-2. Beacon frames and passive scanning were implemented in the first stage. A new timer *BeaconTimer* was implemented to facilitate periodic beacon transmission. A standard C++ list *ap_table* was used to store the RSSI values of the beacons received in all channels and the best AP was selected from the list. For active scanning, probe request and response frames were implemented and the C++ list *ap_table* was used to store the RSSI values. A new timer *ProbeTimer* was created to enable active scanning across all channels, which can be set to *MinChannelTime* and *MaxChannelTime*.

For authentication, authentication frames were created and open-system authentication was implemented, which involves the exchange of two authentication frames. Once a client is authenticated, the AP stores the client's authentication state in a C++ list *client_table*. Association request and response frames were implemented to support association. The AP also stores the client's association state in its *client_table*. Packet filtering and BSS management is done by the AP using *client_table*. The ability to handoff from one BSS to another was also incorporated. The basic IEEE 802.11 handoff mechanism was implemented which involved scanning, authentication and association.

Creating and linking multiple channel objects proved to be very difficult in ns-2, given ns-2's code structure. Hence to create multiple channels without actually creating multiple channel objects, the same channel object was used at different frequencies (channel numbers). Each node as assigned a *Channel_Number* and this is reflected in both the MAC and PHY files of the 802.11 implementation using cross-layer functions. Stations can hear one another only if they share the same channel (channel number). Hence, neighboring BSSs can be setup with its APs assigned different channel numbers creating a multi-channel scenario, without any interference. During scanning, mobile nodes switch their channel numbers using the *CSTimer*, to scan across all channels in the IEEE 802.11 spectrum.

For Inter-AP communication, the distribution system (DS) was implemented using a separate wireless broadcast DS channel for APs. For this, each AP was fitted with an additional network interface and IEEE 802.11 MAC and PHY models dedicated to the DS channel. The DS channel is independent of the channel that the AP uses to communicate with its clients. All our ns-2 enhancements for 802.11 model is available at [14].

To test the FastScan strategy, the newly implemented ns-2 802.11 module was used. The FastScan handoff mechanism was added to the infrastructure model and a C++ standard list was used to implement the databases.

For simulations involving an indoor IEEE 802.11b wireless LAN deployment, a total of 9 APs were deployed in a grid-like pattern. Indoor Access Points are typically deployed with a transmission range of 30 m. Hence, the 9 APs were deployed (as shown in the Figure 5) such that there are over-lapping areas between different BSSs. The APs were deployed such that the neighboring APs and hence the BSSs, operate on non-overlapping channels as much as possible. For IEEE 802.11a, there are 8 non-overlapping channels and the 9 APs were deployed such that the neighboring BSSs all operate on these non-overlapping channels. The number of clients is varied from 1 to 10 in each BSS. Hence a total of 90 mobile stations were deployed across all BSSs.

Table 6 NS-2 IEEE 802.11 MAC/PHY parameters

802.11 Mode	802.11a/b
Maximum Data Rate	11 Mbps for 802.11b 54 Mbps for 802.11a
Center Frequency	2.437 Mhz (Channel 6)
Propagation Model	Shadowing
Receive Threshold	-90dBm
Transmission Power	15 dBm
Beacon Interval	100 ms
MaxChannelTime	11 ms
MinChannelTime	5 ms
Channel Switching Time	5 ms

The ns-2 802.11b simulation parameters are tabulated in Table 6. To simulate VOIP traffic, exponential ON/OFF traffic generators were used over the UDP agents. G.711 codec is used - 64 Kbps coding rate, 20 ms interval, 160 bytes RTP payload size), the parameters of the exponential ON/OFF model are shown in Table 7. This voice traffic generator is capable of handling delays in voice packets of up to 50 ms. This is done by buffering the received packets at the traffic agent layer, so that the stream is smooth even during a handoff. When the transmitter retries the MAC frame for a voice samplem triggering the handoff, the previous samples received by the receiver are buffered at the traffic agent level. Once transmitter's handoff is over, the current sample can then catch up to other samples in the receiver.

Table 7 Voice traffic parameters

Average ON duration	1.000 s
Average OFF duration	1.35 s
Packet size	200 bytes
Sending rate	80 Kbps

Data traffic was simulated using a FTP protocol over a TCP agent. Since it is FTP, there is almost always data traffic in the background and FTP uses a lot of bandwidth due to TCP connections. The parameters of the data FTP model are shown in Table 8.

The mobility model followed by the clients is Random Way Point (the only viasble)and the speed of the client is varied from 1m/s to 10 m/s. In each simulation, the source and destination co-ordinates of the client were randomized in a loop, thereby generating an average of 100 handoffs for random destinations. All the handoff delays were stored and averaged. Wrong handoffs encountered in FastScan (nodes handing off to best APs not in the same direction) led to immediate additional handoffs as corrections. Wrong handoffs and additional correctional handoffs were also included while calculating the average handoff delay.

Table 8 Data traffic parameters

Packet size	1024 bytes
Sending rate	4 Mbps

4.1 Handoff with voice traffic only

In this experiment, only voice traffic was considered. All mobile stations and APs were fitted with VoIP traffic models. Each VoIP connection is a bi-directional UDP flow between the AP and the mobile station. The three components of the handoff delay using the FastScan mechanism is shown in Figure 7.

As the number of voice connections increases, there is a gradual increase in the FastScan handoff delay but it never goes above the 50 ms VoIP threshold. This is due to a gradual increase in channel contention and the resulting backoff. Since voice packets are only sent at 20 ms intervals during the ON period and

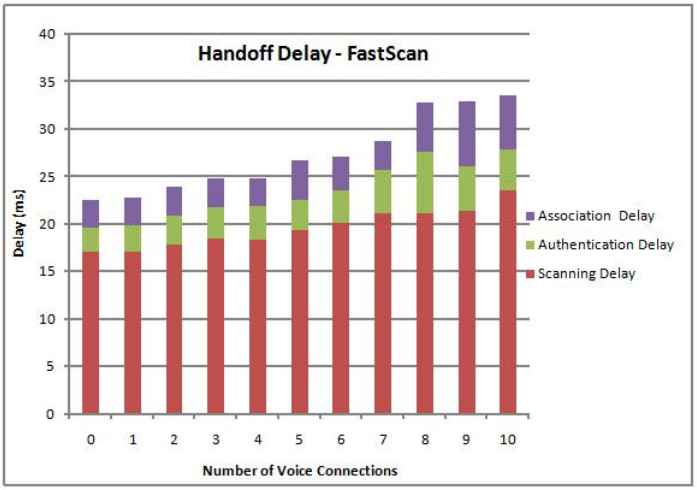

Fig. 7 Split handoff delay of FastScan scheme for varying voice connections

the traffic source does not send any packets during the following longer OFF period, the AP is not fully loaded at any given time instant in the simulation. Hence, the handoff delay does not go above 50 ms even when the number of voice connections is 10.

Though the number of voice connections can be increased further without going over the VoIP threshold, IEEE 802.11b infrastructure mode APs can support only a maximum of 6 G.711 VoIP connections (10 for G.729), even without background traffic. This maximum was calculated and observed with respect to throughput, end-to-end delay and packet dropping probability in [15] and [16]. Increasing it further will harm these other parameters even if the VoIP threshold is not reached. It can also be seen that the scanning delay contributes to almost 90% of the total handoff delay.

In Figure 8, the comparison between FastScan and the basic handoff scheme is plotted. It is observed that the FastScan scheme achieves almost an 80% reduction in handoff latency compared to the basic handoff scheme. The majority of the reduction comes from the scanning phase, where only neighboring channels and best APs are scanned. The basic handoff scheme for IEEE 802.11b deployment using only non-overlapping channels, was also compared against Fastcan. As can be seen, even with only non-overlapping used in active scanning, the handoff delay increased from $41ms$ to $60ms$, averaging around $51ms$, which is still not good enough for voice traffic. This is due to the fact that active scanning still requires the client to stay on channel for atleast *MaxChannelTime*, collecting all probe responses. Figure 9 shows the cumulative distribution of handoff delays for both the FastScan and the basic handoff schemes. It can be seen that the FastScan handoff delay is around $25ms$ while the basic handoff delay is averaged around $140ms$, with $51ms$ for the non-overlapping channels case.

4.2 Handoff with background data traffic

In this experiment, the effect of background traffic on handoff delay is analyzed. At first, 5 voice connections are added in each BSS and then 6 data connections are added one by one.

As can be seen from Figure 10, the first 5 connections are voice and the handoff delays are similar to the previous voice-only results. However, after adding data traffic, the handoff delay shoots up to as high as 44 ms (for 6 data connections). This sudden increase in handoff delay in the presence of data traffic is due to delays incurred by high channel contention, long backoff delays, high bandwidth requirements of the data traffic and the long processing delays at the AP. This is attributed to the continuous traffic load offered by the data connections as opposed to the intermittent traffic offered by the voice connections. Also, since every packet has to be processed and forwarded by the AP, the voice and data performance is limited by the size of the AP buffer. Since the data connections offer a continuous load on the AP, the scanning delay, authentication and association delays all increase due to the AP's limited buffer size, continuous backoffs and high bandwidth required by the data connections.

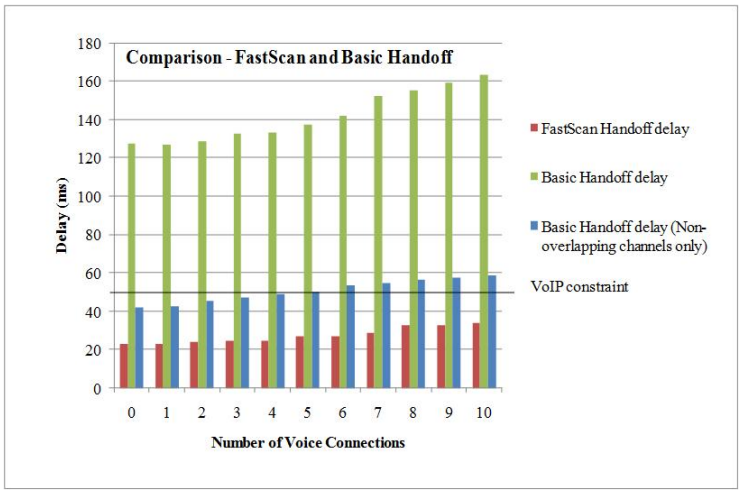

Fig. 8 Handoff Delay comparison of FastScan and basic IEEE 802.11 handoff scheme

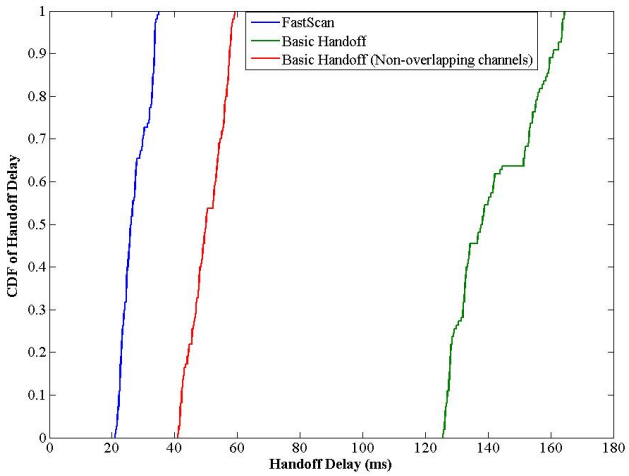

Fig. 9 CDF of handoff delay times

Hence the voice performance of the handoff mechanism is considerably affected by the presence of background data traffic in the network. Studies [15] have also shown that the number of VoIP connections supported by a BSS is also reduced in the presence of data traffic, with respect to throughput and end-to-end delay considerations.

4.3 Effect of handoff on Inter-frame delay

The inter-arrival delay is calculated from the simulation results for a mobile station undergoing a series of handoffs (6 in this case). An IEEE 802.11b scenario is assumed with 9 different BSSs and 5 voice calls in each BSS in addition to the station undergoing handoff. As can be seen in Figure 11, the average inter-frame delay is around 20ms, as the voice packets are sent at 20ms intervals. However, during handoff, the delay shoots up to around 40-45ms because of the handoff delay of 20-25 ms. However the inter-frame delay is still kept under 50 ms.

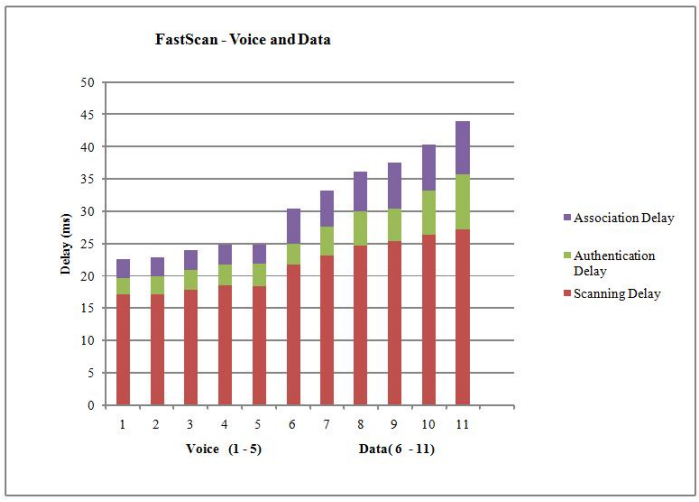

Fig. 10 Split handoff delay of FastScan scheme with background data traffic

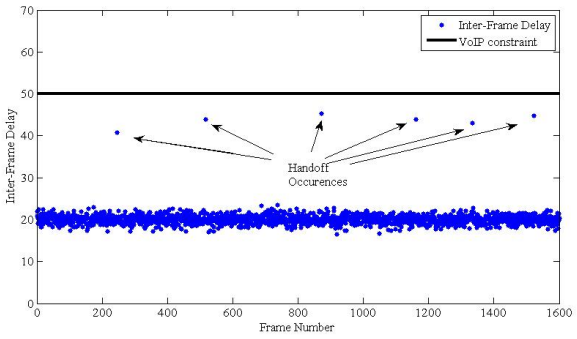

Fig. 11 Effect of handoff on Inter-frame delay

4.4 Wrong Handoffs using FastScan scheme

As previously discussed, wrong handoffs is a problem with the original FastScan scheme, only when database updates do not take place. The same IEEE 802.11b 9-AP deployment in Figure 5 was used and 100 intended handoffs were simulated per load. The traffic load was 5 voice connections plus 5 data connections. Figure 12 shows the average number of wrong handoffs observed for 100 intended handoffs per load, observed in normal scenarios, where the FastScan database is constantly updated by listening to beacons and re-association requests. As can be seen, without the failsafe mechanism in FastScan, the average number of wrong handoffs is usually under 5 and does not increase much with increasing load. In all cases, a wrong handoff was immediately followed by one corrective handoff, resulting in less than 105 total handoffs. With the failsafe mechanism, the number was brought down to 1 or 2 most of the times, by avoiding wrong handoffs if weak probe responses are heard. Thus, in normal operating conditions where the database is constantly updated, wrong handoffs are few in number and the failsafe mechanism helps to further reduce this number. For all loads, the average handoff delay (including wrong handoffs) was under 45 ms, consistent with the observations in Figure 10, both with and without the failsafe mechanism.

In order to better analyze the problem of wrong handoffs and validate the robustness of the failsafe mechanism, a worst case scenario was forcefully simulated. The above 9-AP deployment was used with simultaneous voice and data traffic, but the constant update of database was stopped manually. The mobile clients were allowed to use only the initial databases built during the initial roaming and were not allowed to update them. Without enabling the failsafe mechanism, this worst case scenario resulted in more wrong handoffs, as the best APs either returned very weak probe responses or no probe responses at all (rarely). As shown in Figure 13, the number of wrong handoffs were usually high, around 9 in number. The “no

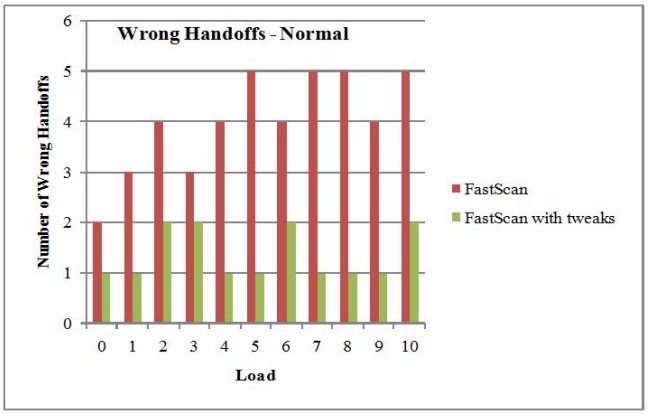

Fig. 12 Wrong Handoffs with FastScan in Normal scenarios

probes heard" scenario happened only once in all the 1100 simulations across all loads, where the client was stuck with its old AP's BSS and the failsafe mechanism kicked in just to handoff the client to the correct AP. By enabling the failsafe mechanism however, most wrong handoffs were avoided by cross-referencing the database and selecting the best AP. The average number of wrong handoffs was reduced to under 3. Even if the database was not updated in this case, the failsafe mechanism was able to do an intelligent hop in its database and pick out the best AP, as the correct best AP's BSS will usually be near to the BSS it is currently leaving. The rare but very severe problem of "no responses heard" was also completely avoided using the failsafe mechanism. Though this worst case scenario rarely happens in normal conditions, the failsafe mechanism is, nonetheless, necessary for the basic FastScan Scheme. Enhanced FastScan, however, does not need this failsafe mechanism, as its location-aware database is very precise and is not completely dependent on constant database update for accuracy. In order to eliminate wrong handoffs entirely, Enhanced FastScan is recommended for IEEE 802.11b scenarios.

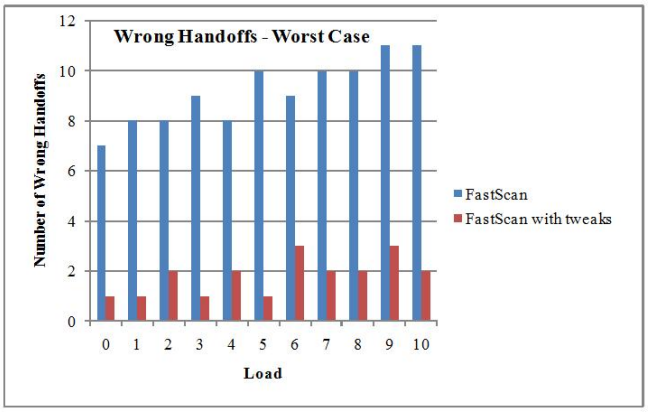

Fig. 13 Wrong Handoffs with FastScan in a worst case scenario

4.5 Enhanced FastScan for IEEE 802.11a deployments

As mentioned earlier, in an IEEE 802.11a deployment, there are 8 non-overlapping channels as opposed to 3 non-overlapping channels in IEEE 802.11b. In this experiment, the dependence of the scanning delay and hence the handoff delay, on the number of non-overlapping channels is analyzed. A mobile station was made to roam across 9 different BSSs operating on non-overlapping channels.

In this simulation experiment, the mobile clients knew the exact location (co-ordinates) of the APs and their own real-time location using X and Y co-ordinates obtained during the simulation to build

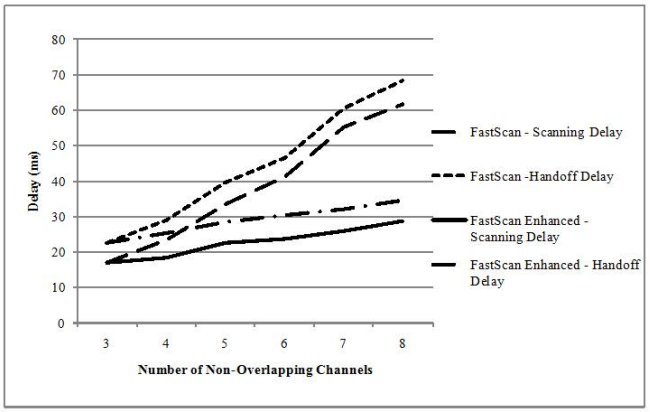

Fig. 14 Scanning and Handoff delays for varying number of non-overlapping channels

the Enhanced FastScan database. For example, the exact AP locations could be available should they be equipped with GPS receivers. If the mobile clients are not GPS equipped, their approximate real-time location can be obtained based on Wi-Fi positioning using triangulation.

From Figure 14, it can be seen that for 3 non-overlapping channels, the FastScan scanning and handoff delays are the same as the IEEE 802.11b results since there are only 2 neighboring channels (and best APs) to scan. As the number of non-overlapping channels increases, the scanning delay is also increased as the number of neighboring channels to scan also increases linearly. For deployments employing over 6 non-overlapping channels, the handoff delay shoots over 50 ms invariably. This is due to the fact that a single current AP in the FastScan database has many neighboring channels (and best APs) entries. The slope for the FastScan handoff delay is observed to be steep. For IEEE 802.11a deployments using 8 non-overlapping channels, the handoff delay is over 65ms, which degrades the performance for VoIP traffic.

When an Enhanced FastScan database is used with direction information, the number of neighboring channels and best APs to scan is reduced depending on the area in which the mobile station is present in the current BSS. In IEEE 802.11a deployments, the number of neighboring channels to scan is only 2-3 (on average) compared to 7 using a normal FastScan database. Hence the scanning and handoff delay is reduced considerably to 34.6 ms, which makes IEEE 802.11a suitable for VoIP traffic. The slope of the Enhanced FastScan handoff delay is much more gradual than the slope of the FastScan handoff curve and is almost flattened, which reduces the handoff delay to around 34 ms. A point to be noted is that these simulation results have been observed in the absence of other traffic in the network; scanning and handoff delays are expected to worsen in actual networks with heavy background traffic.

5 Conclusions

Satisfying QoS requirements of voice traffic over IEEE 802.11 networks, is a formidable challenge because of handoff delays. In this paper, the FastScan scheme is proposed which reduces the scanning delay by using a client-based database. The handoff delay is reduced to as low as $20ms$ for IEEE 802.11b networks, well under the $50ms$ voice threshold. A better scheme Enhanced FastScan, is also proposed which uses the direction and relative position of the client with respect to the current AP. This scheme satisfies voice constraints for IEEE 802.11a networks and provides precise handoffs in IEEE 802.11b networks (avoiding wrong handoffs). To facilitate WLAN simulations for future research in ns-2, a full-featured multi-channel IEEE 802.11 infrastructure mode model was implemented on top of the existing model in ns-2.

These handoff schemes were analyzed using null authentication. If secure authentication schemes like IEEE 802.1X and WEP are used, the authentication delay can last up to a second. Future research in handoff optimization can be directed towards reducing authentication delay in secure IEEE 802.11 networks, using IAPP.

6 Acknowledgements

This work was supported in part by NSF Award CRI CNS 0551686. The authors would like to thank Professor Tom Henderson for his excellent support in implementing the 802.11 infrastructure model in NS-2. The authors would also like to thank the NS developers community for the same.

References

1. IEEE Standard 802.11, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *ANSI/IEEE Std 802.11, 1999 Edition (R2003)*, pp. i–513, 2003.
2. International Telecommunication Union, "General Characteristics of International Telephone Connections and International Telephone Circuits," *ITU-TG.114*, 1988.
3. H. Velayos and G. Karlsson, "Techniques to Reduce IEEE 802.11b MAC Layer Handover Time," *Technical Report TRITA-IMIT-LCN R 03:02*, 2003.
4. Arunesh Mishra, Minho Shin, and William Arbaugh, "An Empirical Analysis of the IEEE 802.11 MAC Layer Handoff Process," *ACM SIGCOMM Computer Communication Review*, vol. 33, no. 2, pp. 93–102, 2003.
5. IEEE 802.11f, "Recommended Practice for Multi-Vendor of Access Point Interoperability via an Inter-Access Point Protocol Across Distribution Systems Supporting IEEE 802.11 Operation," *IEEE Std 802.11F*, 2003.
6. I. Ramani and S. Savage, "SyncScan: Practical Fast Handoff for 802.11 Infrastructure Networks," *INFOCOM, 24th Annual Joint Conference of the IEEE Computer and Communications Societies, Proceedings IEEE*, vol. 1, pp. 675–684, March 2005.
7. Minho Shin, Arunesh Mishra, and William A. Arbaugh, "Improving the Latency of 802.11 Hand-offs Using Neighbor graphs," *MobiSys : Proceedings of the 2nd International Conference on Mobile Systems, Applications, and Services*, pp. 70–83, 2004.
8. V.M. Chintala and Qing-An Zeng, "Novel MAC Layer Handoff Schemes for IEEE 802.11 Wireless LANs," *Wireless Communications and Networking Conference, IEEE*, pp. 4435–4440, March 2007.
9. Vladimir Brik, Arunesh Mishra, and Suman Banerjee, "Eliminating Handoff Latencies in 802.11 WLANs Using Multiple Radios: Applications, Experience, and Evaluation," *IMC : Proceedings of the 5th ACM SIGCOMM Conference on Internet Measurement*, pp. 27–27, 2005.
10. Cisco Press, "Cisco Unified Wireless IP Phone 7920 Design and Deployment Guide," 2005.
11. "Meru Networks," `http://www.merunetworks.com/products/director/director_features.php`.
12. "Extricom," `http://www.extricom.com/content/products`.
13. IEEE Standard 802.11, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 2: Fast Basic Service Set (BSS)," *IEEE Std 802.11r-2008 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008)*, pp. c1–108, 2008.
14. "NS-2 IEEE 802.11 Infrastructure Mode Implementation," `http://ee.washington.edu/research/funlab/ns2_80211.htm`.
15. S. Garg and M. Kappes, "An Experimental Study of Throughput for UDP and VoIP Traffic in IEEE 802.11b Networks," *Wireless Communications and Networking Conference, IEEE*, vol. 3, pp. 1748–1753, March 2003.
16. S. Garg and M. Kappes, "Can I add a VoIP call?," *IEEE International Conference on Communications*, vol. 2, pp. 779–783 vol.2, May 2003.