Enhanced Fast Base Station Switching

Kuang-Hui Chi
Department of Electrical Engineering
National Yunlin University of Science and Technology, Taiwan

J. Morris Chang
Department of Electrical and Computer Engineering
Iowa State University, U.S.A.

Ting-Chung Wang
Institute of Electrical Engineering
National Yunlin University of Science and Technology, Taiwan

Abstract

IEEE 802.16-2009 specifies two fast handover mechanisms, fast base station switching and macro diversity handover, to streamline communication for a mobile station (MS.). Both operate with a diversity set that lists base stations among which an MS can move its connection readily. In view that an unduly chosen diversity set may cause prohibitive cost, we provide means to prevent the diversity set from including base stations that are less likely to serve the MS in the near future. Inspired from the working-set model, our approach develops predictive handover using numerical extrapolation to accommodate temporal locality of the MS. While a joint entry-replacement strategy is exercised to evict least-preferred entries in the diversity set, its capacity is reviewed and tuned periodically allowing for reasonable space demand. Simulation results show that our approach, compared with the counterpart scheme, reduces handover executions by over 48%, handover delay by over 51%, and diversity-set space requirement by up to 17.9% on average. As another salient strength, our approach conforms fully to the standard, keeping current protocols on the MS side operable without modification. Qualitative and quantitative performance discussions indicate the usefulness of our approach in pragmatic settings.

Index Terms: IEEE 802.16, predictive handover, trendline, fast base station switching, macro diversity handover

---

1This work has been supported by the National Science Council, ROC, under grant NSC 97-2219-E-224-001.
2Corresponding author. E-mail: chikh@yuntech.edu.tw. Phone: +885-5-5342601 #4245. FAX: +886-5-5312065.
1 Introduction

Given the prevalence of portable communication devices and a trend toward ubiquitous computing, nowadays a user is enabled to retain network connectivity over wireless media while moving around. Currently IEEE 802.16-based technology is evolving as a new means of public access to a vast variety of Internet services. The coverage or physical constraints of IEEE 802.16 wireless networks, however, may lead a mobile station (MS) to switch its radio link to different base stations (BSs) frequently. Such a handover involves connection teardown and setup, registration and reauthentication with the network, causing a blackout during which the MS cannot deliver data frames to and from the system. Therefore fast handover is essential to streamline communication.

Handover is a process in which an MS migrates from the air interface provided by one BS to another. For this, IEEE 802.16-2009 specifies three mechanisms, namely hard handover, fast BS switching (FBSS), and macro diversity handover (MDHO) [1]. Hard handover is typically termed break-before-make handover, i.e., the MS disconnects from the serving BS before connecting to the target BS. Hard handover is simple but implies abrupt connection transfers across BSs at the expense of longer delay and makes the procedure less reliable [27]. In comparison, FBSS and MDHO cater better for multimedia streaming services like IPTV in a highly mobile environment. These two mechanisms operate with a diversity set, a list of active base stations in range among which the MS can switch its connection seamlessly. While FBSS and MDHO are considered optional for complexity reasons, FBSS remains recommended over hard handover and MDHO to delay-sensitive applications [11].

This paper is concerned with FBSS for its aptness of smoothing over service disruptions. Though effective, the quantitative use of FBSS has not yet been well understood in the literature. A trade-off between benefit and overhead of FBSS relates to diversity-set maintenance. On one hand a diversity set facilitates an MS doing fast handover with a higher success rate, on the other the diversity set should be kept minimal (otherwise activating too many BSs becomes costly.) In this light, we provide means to prevent the diversity set from including base stations prematurely or avoiding base stations that are less likely to serve the MS in the near future. Our treatment founds on numerical extrap-
olation allowing for MS movement behavior. Our objective is to minimize the overhead of FBSS so that it can be made cost-effective to the greatest extent possible.

This study distinguishes itself from previous schemes on IEEE 802.16 fast handover in following aspects:

- Most observed research was focused on reducing hard handover delays and packet loss [6, 9, 10, 12, 17, 23], or conducted with security emphasis, cross-layer considerations [2, 19], or additional network-layer mobility management [7, 14, 16, 18, 25, 29, 36]. Some addressed handover in mesh mode [15] or in a multi-hop relay system [26, 31]. Unlike those work, we concentrate on FBSS; our development is complementary to prior schemes.

- This study provides timely resolution of target BSs for an MS. Although techniques of prospective BS estimation were devised to determine a single target BS [6, 22] or by using Fuzzy logic [3, 21], we avail ourselves of numerical methods to better tolerate errors of estimation due to mobile behavior uncertainty. Consequently our approach is of another utility to observed fast handover schemes. Notably Becvar et al. proposed an effective prediction technique to select the most likely BS [5]. While our approach is an alternative to achieving similar objectives, managing the diversity set in reaction to the predicted dynamic results is indeed our theme.

- As far as compatibility with standard is concerned, previous schemes entail certain modifications to current IEEE 802.16 protocol machinery. In contrast, standard protocols in our architecture remain operable without change. Our proposal does not replace any IEEE 802.16 predefined mechanisms, which characterizes a strength over counterpart schemes.

There is as yet limited research work on FBSS in this area. Among others, Bchini et al. discussed the performance of FBSS for multimedia traffic in high speed mobility scenarios [4]. Wang and Chiang adopted Fuzzy logic to select the best BS for an MS with regard

---

3In [6], an MS is located according to its received wireless link quality. The change of two successively estimated locations gives a movement direction along which a target BS can be resolved, following a similar notion prescribed in [32]. Such assessment brings about nontrivial errors of estimation because of irregular radio signal propagation or multi-path effects. In [22], only a single BS is considered the target to receive handover of the MS. However, mobile behavioral uncertainty may cause fallacious estimation, complicating the handover process contrarily.
to traffic criteria [35]. Fu et al. improved FBSS by means of cell reuse partitioning [13], where some of radio resources of a regular cell were reserved for accommodating handover users with originally poor radio-link performance. Unlike Fu et al.’s scheme that requires modifying standard frame structures, we take a different avenue to enhance FBSS yet preserving the standardized operations. This paper serves a feasibility study on FBSS and MDHO as well. We attempt to corroborate and fit FBSS to practical networks by parameterizing conditions that govern its resulting operational efficiency.

The remainder of this paper is organized as follows. The next section gives a brief background on standardized handover processes. Section 3 elaborates on our proposed methods. Performance evaluation is provided in Section 4. Section 5 concludes this work. Lastly further refinement and pragmatic considerations are covered in appendices.

2 Preliminaries

IEEE 802.16 hard handover involves several stages: cell reselection, handover decision and initiation, synchronization to target BS downlink, ranging, and network reentry. A precursor to cell reselection, network topology information should be known to any interested MS beforehand. Network topology acquisition is achieved by the serving BS periodically broadcasting MOB_NBR-ADV messages, which contain channel information of neighbor BSs and enable the receiving MS to synchronize with neighbor BSs without listening to their DCD (Downlink Channel Descriptor) and UCD (Uplink Channel Descriptor) messages. Therefore, the MS starts scanning neighbor BSs, among which some can be selected as targets. Association may also be carried out so that the MS can acquire information for selecting a proper handover target and/or expediting a future handover to a target BS. For an expository background on hard handover, we refer the reader to [1, Clause 6.3.22.2][28].

FBSS operates using a diversity set for a concerned MS—a list of active BSs with which the MS can move its connection readily, without invoking the lengthy handover process. An FBSS-capable BS broadcasts the DCD message containing the H_Add threshold and H_Delete threshold. These two thresholds serve an MS to decide if MOB_MSHO-REQ should be emitted to request switching to another BS or updating the diversity set. When
the mean carrier-to-interference-and-noise ratio (CINR) of a neighbor BS grows higher than H/Add, the MS sends a MOB_MSHO-REQ message to request adding that neighbor BS into the diversity set. On the contrary, if mean CINR of a BS falls below H/Delete, the MS sends a MOB_MSHO-REQ message to remove that BS from the diversity set.

Figure 1 shows FBSS message flow. Following the reception of MOB_NBR-ADV messages, an MS may send MOB_SCN-REQ to its serving BS and pull a response MOB_SCN-RSP for a time interval for scanning and/or association with a target BS, say BS\textsubscript{J}. The MS may transmit a MOB_SCN-REP message to report its scanning results as dictated by its serving BS. When detecting the average CINR of BS\textsubscript{J} higher than H/Add, the MS sends MOB_MSHO-REQ with a candidate BS set including BS\textsubscript{J}. If BS\textsubscript{J} is also in the BS\textsubscript{I}’s recommended list of MOB_BSHO-RSP, the MS issues MOB_HO-IND \(^4\) for adding BS\textsubscript{J} to the diversity set, thus initiating an update procedure. BS\textsubscript{I} then sends BS\textsubscript{J} HO-

\(^4\)Sending this message may depend on the \textit{Action Time} field of the prior received MOB_BSHO-RSP. Action Time records the number of frames that the BS suggests the MS wait before transmitting a next MOB_MSHO-REQ or MOB_HO-IND.
Request to establish a network connection, and $BS_J$ will generate HO-Response if the establishment is complete. Upon receipt of HO-Response, $BS_I$ sends a MOB_BSHO-RSP message back to the MS for a diversity-set update.

Now that the diversity set accommodates $BS_I$ and $BS_J$, the MS uses MOB_MSHO-REQ to signal a request for switching its anchor BS to $BS_J$, if the CINR of $BS_J$ keeps rising. If the request is granted through MOB_BSHO-RSP, the MS sends MOB_HO-IND to terminate the radio link with $BS_I$ and then performs fast ranging with $BS_J$. Note that $BS_J$ reserves an uplink contention-free ranging subchannel for the MS and places Fast Ranging IE in the extended UIUC (Uplink Interval Usage Code) in a UL-MAP IE to inform the MS of this ranging opportunity. The fast ranging process shall be accomplished in two frames, where the uplink ranging opportunity is indicated by the downlink MAP in the first downlink subframe and then the MS sends RNG-REQ in the successive uplink subframe based on radio parameters recorded in the scanning interval. Subsequently $BS_J$ replies RNG-RSP along with commanded correction information encoded in the second frame. At this point handover is completed.

In FBSS, an MS continuously monitors received signal strength from BSs of the diversity set, does ranging, and maintains a connection identifier with each of them. On the opposite side, involved BSs share or transfer context that includes all the information, particularly authentication state, resulting originally from a network entry procedure. So, an MS authenticated (registered) with any BS of the diversity set behaves as if authenticated (registered) with other BSs of the same diversity set. Notice that every diversity-set update necessitates inter-BS information exchanges over the backbone network, taking some time to establish context at intended BSs. Accordingly the diversity set accounts for cost-effectiveness of FBSS. A proper diversity set enables the MS to hardly migrate off the enclosure of such BSs, implying fast handover to be more likely. Meanwhile, since the diversity set places a burden on the MS and the involved group of BSs, it should be minimized wherever possible.
Figure 2: FBSS performance can be parameterized with several factors. We devise a timely strategy to activate new BSs for FBSS.

3 The Proposed Approach

We shall balance FBSS cost and performance by exercising handover control. To this end, we first provide an overview of our approach (Section 3.1) and a means to select target BSs (Section 3.2.) Considering several factors, Section 3.3 and Appendix A formulate strategies for an MS to activate a new BS for FBSS. In addition to handover control, Sections 3.4 and 3.5 discuss implementation considerations.

3.1 Main Ideas

Our handover control is exemplified in Figure 2, where mobile stations $MS_i$ and $MS_j$ exhibit low and high mobility, respectively. When $MS_i$ moves from base station $BS_I$ toward $BS_J$, we suggest $MS_i$ not including $BS_J$ for FBSS immediately after detecting that CINR of $BS_J$ exceeds $H_{Add}$. This is because premature inclusion of $BS_J$ antedates overhead, which is not necessary though, as $MS_i$ has adequate time to interact for FBSS. On the contrary, $MS_j$ moving faster from $BS_K$ to $BS_J$ should add $BS_J$ in the diversity set upon detecting CINR of $BS_J$ higher than $H_{Add}$. Otherwise $MS_j$ might fail FBSS for losing contact with its anchor $BS_K$ anytime. Accordingly, we instruct an MS to activate new BSs for FBSS neither too early nor too late in reaction to mobility rate. This is done...
by the MS starting the diversity-set update procedure at suitable time, advised by its anchor BS sending MOB_BSHO-RSP with a given Action Time. Further, provided that the diversity set is limited to a few entries for cost reasons, we use numerical extrapolation to keep BSs in the set that are most likely to serve the MS in the near future.

Moreover, as evidenced in substantial literature [8, 33], user mobility patterns feature locality, a property common to user behavior in nature. This text restricts attention to temporal locality, i.e., a BS visited by an MS tends to be revisited in the near future. Figure 2 also depicts temporal locality of MSs that roam to and fro between BSs. Temporal locality of movements is common in daily life. For example, a user mostly migrating among certain workplaces on a regular basis. In view of temporal locality, we shall keep the diversity set meeting with MS’s movement patterns in a reasonably affordable way. This is done by capturing per-MS locality under the notion of working set used in many studies on program paging behavior. By doing so, the diversity set is becoming stabilized for some time or evolving with gradual updates along with transitional changes in movement patterns. Since the diversity set is managed in anticipation of FBSS operations, the MS can still perform fast transitions among BSs in a cost-saving fashion.

### 3.2 Selection of Target BSs

Let us first deal with target BS selection, prerequisite for any handover process. When an MS detects that CINR from its current BS has fallen below some threshold, the MS performs neighbor BS scanning in light of the MOB_NBR-ADV message. Scanning results give example CINRs as shown in Figure 3(a). Suppose that detected CINRs by the MS for one neighbor BS, $BS_J$, in this case, and its current BS at different time $t_i$ are plotted in Figure 3(b). Viewing each CINR value as a data point, we have a sequence $(t_0, y_0), \cdots, (t_4, y_4)$. Given $n$ such data points, it is apropos to derive a best-fit function $\psi$ that makes a trendline reflecting how CINRs from $BS_J$ vary. We adopt a logarithmic trendline in the form of $\psi(t) = a \ln t + b$, where

$$a = \frac{n(\sum t_i y_i) - (\sum t_i) (\sum y_i)}{n(\sum t_i^2) - (\sum t_i)^2} \quad \text{and} \quad b = \frac{\sum y_i}{n} - a \frac{\sum t_i}{n}.$$ 

Similar assessments result in another trendline for $BS_K$. Comparing such derived trendlines leads us to infer that the MS is more likely to migrate to $BS_J$ than to $BS_K$. 

8
CINRs from BSs tend to increase CINRs for the MS during latest measurements:
From BS_I: 20.5, 18.6, 12.2, 15.6, 13.5 (dB);
From BS_J: 6.8, 10.7, 20.2, 18.5, 16.1;
From BS_K: 6.5, 8.3, 7.8, 7.5, 6.8

Figure 3: Changes of CINR for an MS with respect to neighbor BSs suggest toward which BSs the MS tends to migrate.

Potential target BSs are evaluated in the above fashion. For tolerating mobile direction uncertainty, an MS may select most likely k out of neighbor BSs with average CINR higher than H_Add to form a set S that indicates handover targets. (However, if such selection is not available due to falling short of H_Add, the MS proceeds to the conventional hard handover process.) If S is a subset of the MS’s diversity set, then the MS is migrating within a region encompassed in its current diversity set. In this scenario, normal FBSS operations satisfy; no further measures need to be taken. On the other hand, if S contains some objects new to the diversity set, S identifies eligible, new BSs for FBSS but the MS may defer activating them as per the next subsection.

3.3 Predictive Handover

We now clarify when to activate a new BS in S, by finding the future latest point for an MS to update the diversity set to accommodate the new BS. Although the MS normally persists with its current BS, for connectedness the MS needs to foresee when the radio link from the BS may deteriorate to an extent that its received bandwidth (quality of service) cannot last any longer. Before that point, FBSS should commence. Such foreknowledge is vital if the MS will actually migrate to a new BS in S.
Since the achievable bandwidth is subject to CINR, we denote by $X$ the least CINR level that sustains the minimum bandwidth requirement by the MS. ($X$ is typically known a priori.) Subsequent to resolving a trendline $\psi$ for the current BS, the farthest time instant $\hat{x}$ for $\psi(\hat{x}) \geq X$ can be found accordingly. Figure 3(b), for instance, shows the deduced point $\hat{x}$ and a forthcoming handover thereafter. Such foresight benefits the MS to start activating $BS_J$ in $S$ for FBSS as time approaches $\hat{x}$, if some hysteresis margin of CINR between BSs for avoiding ping-pong effects is further ensured. To realize, given that the mean $\bar{t}$ and standard deviation $\sigma$ of time required for completing diversity-set updates are known, we suggest that the MS initiate the update procedure no later than the point $(\hat{x} - \bar{t} - 2\sigma)$, i.e., advancing the start of the procedure by $(\bar{t} + 2\sigma)$ before the deadline $\hat{x}$ according to the 68-95-99.7 rule [34]. The rule implies that for a normal distribution, about 95% of diversity-set updates finish within time interval $[\bar{t} - 2\sigma, \bar{t} + 2\sigma]$ in length. Therefore, if the MS triggers a diversity-set update at the suggested time, the probability of completing the procedure after $\hat{x}$ will approximately be 0.025 (half 5%), meaning a high likelihood of FBSS continuity.

Following handover preparation is handover execution that takes trendlines into account. That is, the MS may opt to switch its anchor BS to the one with highest projected future CINR (along its trendline) that surpasses the current BS. By doing so, the MS can maintain its connectivity with some BS capable of provisioning higher bandwidth. In other words, our approach enhances FBSS in that BS switching allows for moving trends of CINR changes over some time with projections. This reduces ping-pong effects when the MS transits among BSs.

Our trendline has implicitly subsumed the factor of per-MS mobility rate because it indicates how fast CINRs with respect to a BS vary. (The faster the MS moves, the more the trendline will slope.) We exploit a logarithmic trendline to tolerate temporary fluctuations in CINR out of multi-path effects or other causes. In line with overall trend, our prediction thus becomes less sensitive to occasional variations. A small number $n$, say 5, of CINR readings would suffice. It is also practical to designate the BS to calculate trendlines collectively, as shall be clarified next.

---

5IEEE 802.16 specifies normalized CINR per modulation indicative of data rate in Clause 8.3.7.4.2 [1].
6These statistics relate to network dynamics of inter-BS information exchanges for context establishment across the backbone.
### 3.4 Refinement

For forecast timing of FBSS activities, all necessary computations can be shifted to the BS side so that a BS performs estimations on behalf of local MSs. This is done by an MS transmitting MOB_SCN-REP messages to report its collected CINR metrics about neighbor BSs, whenever available, to its anchor BS. This enables the BS to learn MS-perceived CINR information. Sometime later, due to an inclination to initiate FBSS, the MS reports CINRs of target BSs using a MOB_MSHO-REQ message. In response, the anchor BS settles the point \( \hat{x} - \bar{t} - 2\sigma \) as per Section 3.3, whose value is then carried in the Action Time field of a MOB_BSHO-RSP message addressed back to the MS. In this manner, the MS functions transparently to our proposed operations; current protocols on the MS remain operable without alteration. Another merit is that our design makes the MS knowledgeable about when to trigger diversity-set updates, without undergoing any expensive computations.

If the diversity set is limited to a small number \( k \) of BSs, \( \mathcal{S} \) is well suited to serving as the diversity set that includes the most likely next BSs. Still the diversity set shall evolve with gradual changes as the MS migrates with spatial continuity. In our architecture we let a new BS in \( \mathcal{S} \) replace a BS of the current diversity set with lowest projected future CINR. Such a replacement is called a dropout in this text, meaning that a BS remaining qualified to belong in the diversity set has forcibly been edged off the set for space reasons.

### 3.5 Diversity-Set Maintenance

As mentioned, an MS probably exhibits temporal locality. In order to capture locality, we use a parameter \( \Delta t \), similar to the working-set window, for keeping track of which BSs were visited during the latest period of \( \Delta t \) time units. These BSs may form a diversity set. Figure 2 illustrates that sometime MS\(_i\) assumed a diversity set \( \{BS_I, BS_J\} \), whereas MS\(_j\) a set \( \{BS_K, BS_I, BS_J\} \). Here letting \( \mathcal{S} \) assume the diversity set, the MS can determine its demand for diversity-set size adaptively every \( \Delta t \) time units. This is done by comparing two statistics timeframe by timeframe. One is the ratio \( \varrho \) of dropouts to the total number of handovers occurred in the most recent timeframe; \( \varrho \) signifies space deficiency of diversity set. The other is the ratio \( \eta \) of \((|\mathcal{S}|-1)\) to 1 less than the average
number of the MS’s neighbor BSs; $\eta$ represents space sufficiency of diversity set. These two statistics can give an indication of current status of the diversity-set size, i.e., $\varrho > \eta$ meaning that space deficiency outweighs space sufficiency and that the diversity set should be enlarged. Conversely, a contraction of the diversity set is allowable. Subtracting 1 from both the numerator and denominator of $\eta$ is to exclude the current BS from potential next targets, because the current BS is likely to remain in the diversity set but handover from the current BS to itself hardly occurs. Neighbor BSs in range are counted whenever the MS performs scanning, from which an average value of neighbor BSs during the last timeframe can be obtained.

Our diversity-set maintenance algorithm is summarized in Figure 4 (lines 19 to 24), where aforementioned operations undertaken by an MS are also recapped for comprehen-

```plaintext
var k: integer; // (input) diversity-set size, initially 2
S: set of integer; // (output) most-likely BSs as the diversity set
begin
  On detecting CINR from the current BS below a threshold:
  i := 0; // counter for scanning
  set a timer $\delta$; // timeout triggers neighbor BS scanning
  On timeout of $\delta$: // cell reselection, handover initiation
  scan BSs within range into a set $D$; // set of neighbor BSs
  forall s $\in D$ do $y[s] :=$ CINR of s; // array $y$ is indexed by s
  send MOB_SCN-REP($D, y$) to the current BS; // trendline data
  i := i + 1;
  if $(i < n)$ then set a timer $\delta$ // to schedule the next scan
  else
    begin
    send MOB_MSHO-REQ($D, y$) to the current BS;
    receive MOB_BSHO-RSP($D'$) from the current BS
    // $D'$ is an orderly array of recommended target BSs
    end
  On timeout of $\Delta t$: // time for diversity-set dimensioning
  compute $\varrho$, $\eta$;
  if $(\varrho > \eta)$ then $k := k + 1$ // adjust diversity-set size
  else if $(\varrho < \eta)$ then $k := k - 1$;
  $S := \{\}$;
  for $l = 1$ to $k$ do $S := S \cup \{D'[l]\}$ // top-ranking $k$ BSs of $D'$ form $S$
end
```

Figure 4: A diversity-set maintenance algorithm performed by the MS. The algorithm is event driven, relying on two timers for scheduled operations.
siveness. Diversity-set capacity is reviewed on each timeout of $\Delta t$, at the end of each timeframe. Examining the condition ($\varrho > \eta$) gives an indication of whether the diversity-set size $k$ should be increased. If the condition holds, the space allocated to the diversity set appears smaller than required, so $k$ is enlarged. On the other hand, $k$ is diminished if $\varrho < \eta$. Subsequently the MS singles out first $k$ BSs from an orderly array $D'$ of potential next targets. $D'$ is received via a MOB_BSHO-RSP message from the current BS (line 16) that is charged with calculating trendlines of CINRs for BSs in radio range of the MS as in Section 3.4. Based on CINRs measured at times $t_0, t_1, \cdots, t_n$ (with interspace $\delta$) and the resultant trendline $\psi$ for each neighbor BS, the current BS assigns in $D'$ identities of neighbor BSs in descending order of their produced $\psi(t_n + 3\delta)$. After acquiring $D'$, it is within the MS’s discretion to draw top-ranking $k$ BSs from $D'$ to form its diversity set.

In passing, lines 4 to 6 of Figure 4 prepare for cell reselection, and lines 9 to 18 describe $n$ iterative scanning ahead of handover initiation. Triggered by a $\delta$-timer expiration, each iteration begins with discovering a neighbor BS set $D$ and recording CINR of every such BS in an array $y$. Then the MS transmits a MOB_SCN-REP message containing $D$ and $y$ whereby the BS side can accumulate data for trendline calculation. Such an interaction repeats $n$ times till the MS sends MOB_MSHO-REQ to its serving BS to signal the need to compute trendlines. Therefore the BS derives trendlines in question over which future CINRs of BSs in $D$ are projected, so as to recommend $D'$ to the requesting MS. Indices of $D'$ represent the preferred order of potential next BSs for the MS.

4 Performance Discussions

Our approach is compared with the base FBSS scheme through simulations. The base FBSS scheme refers to any fast handover schemes including [1, 13] that manage diversity sets based on standard H_Add and H_Delete without additional intervention.

4.1 Experimental Environment

We developed a custom event-driven multithreaded simulator in C that mimicked an MS moving within an IEEE 802.16 network in a discrete time-based fashion (Figure 5.)\footnote{The program is available at URL http://www.ee.yuntech.edu.tw/~winlab/EnhancedFBSS.zip.} The
Figure 5: A network of 9 equally-spaced BSs with an identical communication range \( r \). Starting from near the center, the MS was simulated to move about at a constant speed. The field covered by the network was viewed to consist of \( 4 \times 4 \) virtual square grids whose sides had length \( d \). Situated at each grid intersection was a BS with spherical radio range \( r \). Since IEEE 802.16 was designed mainly for metropolitan access, we adopted the Manhattan mobility model yet with minor variants, by confining the MS to move in one of four mutually perpendicular directions. The MS migrated at 5 or 72 km/hr (assuming a pedestrian or vehicle) whose entire trajectory was composed of consecutive short courses. Every course began by the MS deciding to head to one direction and continued till it moved rectilinearly along the chosen direction for 1 or 30 time units. Where a course ended, another new course regenerated straightway but its heading direction was chosen anew from the four equiprobable directions. If the MS was about to migrate off the field, its next coordinate was still bounded by the borderline, resulting in a movement along the boundary sometime. Observe that our pedestrian scenario resembles a Random waypoint model in sense that the MS movement direction changes every time unit.

Allowing for that CINR dropping below a threshold will trigger the scanning procedure, we let the MS scan neighbor BSs when the received signal strength (RSS) from its serving BS deteriorated below \(-80 \) dBm. (Note that, for simplicity, RSS was adopted in lieu of CINR throughout our experiments to measure radio channel conditions.) We varied \( r \) in terms of \( d \) and the combination of \( H_{Add} \) and \( H_{Delete} \) to achieve different cell coverage of BSs so that a diversity set comprised a different number of qualified BSs. Performance results were collected assuming \( r \) to be \( 0.8d \), \( d \), and \( 1.2d \), respectively, while the
Table 1: Parameters used for simulation.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation time</td>
<td>3600 (time units)</td>
</tr>
<tr>
<td>Warm-up time</td>
<td>120 (time units)</td>
</tr>
<tr>
<td>Distance between nearest BSs (d)</td>
<td>1 (km)</td>
</tr>
<tr>
<td>Per-BS communication range (r)</td>
<td>0.8d, d, 1.2d (km)</td>
</tr>
<tr>
<td>BS maximum transmit power</td>
<td>43 (dBm)</td>
</tr>
<tr>
<td>Height of each BS (h_{BS})</td>
<td>30 (meters)</td>
</tr>
<tr>
<td>Height of the MS (h_{MS})</td>
<td>1.7 (meters)</td>
</tr>
<tr>
<td>Shadowing effect</td>
<td>[−8, 8] (dBm)</td>
</tr>
<tr>
<td>Hysteresis margin</td>
<td>1 (dB)</td>
</tr>
<tr>
<td>IEEE 802.16 frame</td>
<td>20 (ms)</td>
</tr>
<tr>
<td>Threshold to trigger scanning</td>
<td>−80 (dBm)</td>
</tr>
<tr>
<td>⟨H_{Add}, H_{Delete}⟩</td>
<td>⟨−86, −84⟩ or ⟨−84, −82⟩</td>
</tr>
<tr>
<td>Max number of scanning (n)</td>
<td>5</td>
</tr>
<tr>
<td>Inter-scanning interval (δ)</td>
<td>50 frames</td>
</tr>
<tr>
<td>Scanning interval</td>
<td>1 (time unit)</td>
</tr>
<tr>
<td>Diversity-set update</td>
<td>1 (time unit)</td>
</tr>
<tr>
<td>Window ∆t</td>
<td>360 (time units)</td>
</tr>
</tbody>
</table>

combination of H_{Add} and H_{Delete} was represented by tuple ⟨−86, −84⟩ or ⟨−84, −82⟩ (in dBm) for short hereinafter. The tuples signify two configurations pertinent to system planning and management. The three cases, r = 0.8d, r = d, and r = 1.2d, represent different BS density such as in suburban or urban areas.

Table 1 lists system parameters in use. According to the Okumura-Hata model [30], the path loss along a terrestrial radio link takes $46.3 + 33.9 \log(f) - 13.82 \log(h_{BS}) - [3.2(\log(11.75 \cdot h_{MS}))^2 - 4.97] + [44.9 - 6.55 \log(h_{BS})] \log(r')$ in metropolitan areas, where $f$ denotes the transmission frequency (measured in MHz, instantiated to 3500 here), $h_{BS}$ and $h_{MS}$ are heights of the BS and the MS (in meters), respectively, and $r'$ is the distance between the BS and the MS (in km.) Supposing that the BS maximum transmit power was 43 dBm, the propagation path loss amounted to 135.648 + 35.33486 log($r'$) in our simulation. Besides, a shadowing effect causing another possible loss of −8 to +8 dBm was taken into account. For a given set of parameters, 10 independent identical simulations were conducted on base FBSS and our scheme, respectively, to produce average performance statistics.
Figure 6: Comparison of handover counts for a slow MS. In the same circumstance our approach improves by over 87%. In (b) and (c), hard handovers and fast handovers of our scheme under $\langle -86, -84 \rangle$ become so rare as to be rounded to zero.

Figure 7: Comparison of handover counts for a fast MS. In the same circumstance our approach improves by over 48% (up to 77.1%).

### 4.2 Handover Count

The first performance index is concerned with how many handovers were performed during our simulation. Figures 6 and 7 compare average handover counts of the base FBSS scheme with those of our approach in different network coverage. In these figures, the lower part of each bar represents hard handover, whereas the upper part fast handover thanks to FBSS. Note that costly hard handover should be avoided wherever possible, so the occurrence of hard handover matters. Figure 6(a) shows that our scheme reduces hard handovers by around 93.3%. Figure 6(b) indicates that, if the communication range $r$ reaches $d$, our scheme reduces hard handovers by 87% and nearly 100% when $\langle H_{Add},$
\(H_{\text{Delete}}\) equals \((-84, -82)\) and \((-86, -84)\), respectively. Figure 6(c) reveals that our approach prevents unnecessary handovers as long as network coverage permits.

As to a fast MS, Figure 7(a) shows that our scheme reduces hard handovers by 70.1\% and 77.1\% under the two \(\langle H_{\text{Add}}, H_{\text{Delete}} \rangle\) conditions. Figure 7(b) implies that our reduction of hard handovers amounts to 67.8\% and 85.7\%, respectively. Figure 7(c) shows that our approach makes a saving of hard handovers by 59.3\% and 48\%. Apart from hard handover, our approach also helps an MS moving at a vehicular speed decrease fast handover execution by an appreciable amount. On the whole, Figures 6 and 7 show a marked improvement of our approach over the base FBSS scheme.

Our performance gain stems from our predictive handover mechanism. Recall that our approach employs trendlines and numerical extrapolation to resolve the most likely BSs for the MS, as opposed to the conventional FBSS scheme that performs handover based on mean CINRs. In comparison, we utilize the changes of CINR from candidate BSs to decide which potential next BSs can serve the MS best. Such a decision is made alongside whether the target BS still outscores the current BS in terms of the projected future CINR. If so, handover is carried out. Our design prevents the MS from switching over to BSs that are momentarily providing high CINR but channel qualities soon deteriorate due to user’s mobility or time-varying channel conditions. Therefore, either hard handover or fast handover to such provisional BSs becomes less probable in our architecture.

Moreover, a notable finding from Figure 6 is that, although a lower \(\langle H_{\text{Add}}, H_{\text{Delete}} \rangle\) may appear to invite more potential next BSs, a randomly walking MS does not necessarily benefit from the standard FBSS. To see this, let us restrict attention to the base FBSS part of Figure 6(a), which suggests that an MS in \((-86, -84)\) contexts encounters more handovers. This is because, for a random walker who changes heading direction all the time, a lower \(\langle H_{\text{Add}}, H_{\text{Delete}} \rangle\) also implies a higher likelihood that a target BS drops out of the diversity set, if movement away from that BS happens. When the diversity set shrinks and contains single BS (namely anchor BS itself), the MS can only invoke hard handover at that moment. Put another way, a higher \(\langle H_{\text{Add}}, H_{\text{Delete}} \rangle\) means a stabler diversity set, i.e., adds and drops to the diversity set become less frequent, exposing the MS to less chance of performing hard handovers at the incorrect time. Such argument holds for Figure 6(b) as well, where coverage of BSs has better overlap. However, when
coverage overlaps to a sufficient extent as reflected in Figure 6(c), the diversity set becomes more likely to accommodate multiple BSs often. As a result, the growing number of handovers are offset to some degree by then.

The value of reducing handover executions is multifold. First, the MS is favored with smooth communication activities and is rid of putting much effort into frequent handover handling. This advantages the network side in the meantime. For instance, BSs are thus relieved of excessive radio link setup and teardown, registration and reauthentication processing, and internal resource management, to name a few. Second, when it comes to hard handover or fast handover count, Figures 6 and 7 show that our approach appears relatively insensitive to different $\langle \text{H}\_\text{Add}, \text{H}\_\text{Delete} \rangle$ configurations. In particular, the number of invoking hard handover in all the subfigures does not differ materially. So does fast handover, compared with the base FBSS scheme. This finding justifies another merit of our approach, suggesting that our design contributes to stable outperformance in whatever network settings.

### 4.3 Handover Delay

The statistics of Figures 6 and 7 serve a purpose of obtaining handover delay experienced by the MS. In this regard, we examine how the two subject schemes perform in the presence of hard handover and fast handover taking on distinct time characteristics. For simplicity, let $w$ denote the weight of hard handover delay relative to fast handover delay. Supposing that fast handover delay is unified to 1 on average, the total handover delay the MS underwent during simulation consists of fast handover count and the product of $w$ and hard handover count. Following such assessments in a given $\langle \text{H}\_\text{Add}, \text{H}\_\text{Delete} \rangle$ setting yields two quantities, one for the base FBSS and another for our approach. The proportion of the difference between the two quantities (reduction by our approach) to total handover delay incurred by the base FBSS scheme gives a second metrics. Our outperformance in this dimension is depicted in Figure 8. Subfigures (a)(b) plot results for our pedestrian scenario, and (c)(d) for our vehicle scenario.

Overall, Figure 8 shows that our approach saves handover delay by over 51%, while the saving shall become stabilized when $w$ grows large. Such steadiness is attributed to that a larger $w$ makes hard handover a dominant fraction of total handover delay, whence
Figure 8: Handover delay reduction versus weight $w$. Our approach improves by over 51%. Subfigures (a) and (b) are data plots for a slow MS, and (c) and (d) for a fast MS. Note a scale shift in the vertical axis for better illustration.
the metrics in proportion form has $w$ canceled out (both numerator and denominator have $w$ in common), turning out to be some constant eventually. In general, our scheme produces more evident outperformance in the pedestrian scenario in that our reduction of hard handovers is apparent on a scale of several tens of handovers (Figure 6.) Such reduction would appear less significant when handovers totals several hundreds (Figure 7.)

Figure 8 indicates that handover delay reduction may augment when more cell coverage overlaps. However, the relative outperformance in the context of $r = 1.2d$ varies greatly between different mobility models. For reasoning, observe that the condition $r = 1.2d$ corresponds to an environment with densely overlapping cell coverage in which the MS can readily be serviced by multiple BSs all the time. Such network settings are in favor of FBSS, causing fewer occurrences of hard handover and thus making our performance gain comparatively less prominent in contrast to where BSs have smaller communication range. Again, when the reduction of hard handovers is related to the total number of handovers in pedestrian or vehicle scenarios, we have different outperformance results.

Considering a fast MS, Figure 8(d) shows that our scheme may achieve a reduction of handover delay slightly less than that in Figure 8(c). This is because in a network setting with $(-86, -84)$, the diversity set may admit more BSs (see also Figure 9). In that case, FBSS is more probable; heavyweight hard handovers occur less often. As hard handovers become fewer, our approach plays a lesser role in reducing handover delay and therefore brings a slightly lesser reduction. Nonetheless, it is worth stressing that different $\langle H_{Add}, H_{Delete} \rangle$ configurations have dissimilar utilities. For example, cross-examining Figures 6, 7, and 8 give a subtlety that the $(-86, -84)$ configuration expedites handover processes particularly in handover delay reduction, while giving rise to hard handover occurrences. The subtlety suggests that no single $\langle H_{Add}, H_{Delete} \rangle$ serves all interests. A higher $H_{Add}$ meets a demand for fewer handover executions and shorter handover delay, whereas a lower $H_{Add}$ for more significant diminution in handover delay in a relative sense.

### 4.4 Diversity-Set Size

With regard to how many BSs are involved for mobility support, let us take the fast MS scenario as an example. Figure 9 compares average diversity-set sizes of the two subject
Figure 9: Comparison of average diversity-set sizes, number of active BSs for servicing a single BS. Under the same condition our approach improves by up to 17.9%.

schemes over the simulation period. The figure demonstrates that our approach keeps the diversity set compact. More specifically, subfigure (a) shows that, given \( r = 0.8d \), our scheme diminishes the diversity set by 16.3% and 8.8% when \( \langle H_{\text{Add}}, H_{\text{Delete}} \rangle \) is configured to \( \langle -84, -82 \rangle \) and \( \langle -86, -84 \rangle \), respectively. Subfigure (b) shows that our scheme yields a saving by 17.9% and 5%, respectively, under the two \( \langle H_{\text{Add}}, H_{\text{Delete}} \rangle \) conditions. Subfigure (c) reflects that, for \( r = 1.2d \), our approach reduces the diversity set by -3% and 9.4%, respectively. These subfigures assure that our approach maintains the diversity set of less than 2 BSs most of the time (see also Figure 11.) Since the diversity set always contains at least one entry, i.e., the MS’s current BS, our approach is effectual for retaining at most another BS in the diversity set even if a plurality of BSs are qualified for FBSS operations.

As a remark, Figure 9(c) indicates that our approach might lose its advantage sometime. This leads us to further investigate how the occupancy of the diversity set changes with mobility randomness. Letting the MS move at 72 km/hr throughout but select its movement direction after a period of certain time units repeatedly, the interval between successive selections of heading direction defines a *mobility epoch* here. Figure 10 shows that our approach may cause a broader diversity set in some cases. An underlying reason is that, as stated in Section 3.5, we adjust diversity-set capacity adaptively every \( \Delta t \) time units. Hence, there exists some possibility that an obsolete BS resides in our diversity set before capacity adjustment is done upon timeout of \( \Delta t \).
Figure 10: Average diversity-set size versus mobility epoch length ($r = 1.2d$).

Figure 11: Comparison of diversity-set sizes along time line ($r = d$).

Notice that Figure 9 shows time-average statistics. Although our approach may only produce a saving of several percents in this aspect, the resulting profits remain pronounced if the length of an observation period is taken into account. Considering our simulation with an observation period of several thousands time units, cumulative profit gains over the entire period are indeed substantial. Our efficacy of minimizing the diversity results from: 1) employing trendline and extrapolation techniques to determine prospective BSs among potential targets, and 2) moderating diversity-set capacity adaptively on a periodical basis by assessing space deficiency ($\varrho$) against sufficiency ($\eta$). The latter measures are crucial to restrain the diversity set from growing larger than necessary.

To demonstrate other further information than time-average values, Figure 11 shows
how the diversity set varies in size over time at any point during simulation. For concise presentation, we confine ourselves to the case of \( r = d \) as an example. (Performance statistics in other settings can be estimated proportionately in scale with reference to Figure 9.) From Figure 11 it can be seen that our approach is effective in preventing more than 2 BSs from being added in the diversity set. Besides, compared with the counterpart scheme, our diversity set encounters smaller-scale fluctuations in size, free from growing or shrinking rapidly. Such lower variance can lessen the overhead of diversity-set space management in a sense. In addition, cross referencing Figures 11(a) and (b) ensures our approach to behave moderately despite different \((H_{\text{Add}}, H_{\text{Delete}})\) configurations. This implies stable outperformance in diversity-set dimensioning.

To conclude this section, we remark that our approach represents a cost-effective design, trading pre-handover computation for FBSS efficiency in terms of practical handover count, handover delay, and diversity-set capacity. While maintaining a diversity set of a proper size is a trade-off issue, Figures 6 to 10 establish that this issue has been dealt with in our architecture because our approach not only improves handover performance but also reduces resources requirements for the diversity set. We thereby lighten a burden on both the MS and BS for FBSS operations. Meanwhile, in system perspective, our approach entails a smaller group of BSs servicing an MS. Unaffected BSs are thus capable of sparing other local MSs more air time for communication. In consequence, overall network throughput and channel availability also increase. Characterizing potency in several important aspects, our treatment can make FBSS better accessible to mobile users as well as system operators.

5 Conclusion

This paper presented means to ameliorate FBSS for mobility support in IEEE 802.16-based networks. Mobility support approached by FBSS warrants closer study. We tackled the problem in question by using regression and numerical methods to project the future behavior of a concerned MS. Considering further that an MS tends to exhibit temporal locality, we leveraged the working-set model to capture locality peculiarity. The captured locality was then taken in diversity-set management. In addition, we introduced a joint entry-replacement strategy to evict least-preferred entries in the diversity set, if
the diversity set incorporating new BSs overran its tentative space limit. Entry eviction was done according to aforementioned projected results, while the space limit was tuned periodically to allow for reasonable diversity-set space demand. Such space demand was evaluated whenever a timeframe of $\Delta t$ time units had elapsed, leading to subsequent increase or decrease of the diversity set by 1 in capacity. The new space limit was hereon put into effect throughout the following timeframe.

In our architecture, the diversity set was evolving reactively to changes in MS’s movement patterns. Our diversity set was maintained in anticipation of FBSS machinery, so the MS was enabled to carry out fast handover in an inexpensive manner. This was due to a refinement of our design shifting all necessary computations to the BS side (Section 3.4.) Another reason was that our design made the MS knowledgeable about when to trigger diversity-set updates without undergoing costly computations (Section 3.5.) Our approach was developed for ease of application by a vast number of end-user devices in that a BS is charged with performing most computations on behalf of local MSs; current protocols on the MS do not require tailoring. Apart from backward compatibility, these traits also imply interoperability with other standardized protocols elsewhere, meaning ample usefulness of our design. Quantitative discussions of Section 4 showed that our approach avoided invoking handovers, diminished handover delay, and alleviated the overhead of managing diversity sets by an appreciable amount. These strengths indicate that our approach would get fielded promiscuously in practice.

In closing, we outline two future directions to work on. First, our approach lends itself to an ingredient of cross-layer handover schemes. Cross-layer fast handover generally relies on using link-level events to trigger higher-level handover of an MS. Such schemes benefit users moving across different administrative domains. Another direction is to incorporate our design in heterogeneous networking as per IEEE 802.21. In an IEEE 802.21 framework, an information server is deployed to maintain network topology of an autonomous system. The server is of utility for an MS to learn information like neighbor maps or link-level parameters, loading status, and available services of other BSs which is inaccessible through typical neighbor BS scanning. In some critical situations, e.g., an ambulance likened to the MS speeding to a hospital, the server may be tasked to intervene in resources reservation at BSs en route ahead that are not yet within radio range.
of the MS. In this respect, locality shall be managed \textit{proactively} rather than reactively, in conjunction with our predictive handover mechanism. It is practical to extend our development to where original FBSS is not yet operable. The server may also be designated to assist resolving prospective BSs timely for the MS about to perform vertical handover. These research directions are worth addressing in another monograph.

\section*{References}


Biography

**Kuang-Hui Chi** received the Ph.D. degree from National Chiao Tung University, Taiwan, in 2001. He is an associate professor at National Yunlin University of Science and Technology, Taiwan. He was with the Computer and Communications Research Laboratories, Industrial Technology Research Institute, ROC. His current research interests include Wireless Internet and Protocol Verification. He is a member of the ACM and the IEEE.

**J. Morris Chang** received the Ph.D. degree from North Carolina State University. He is an associate professor at Iowa State University. His industrial experience includes Texas Instruments, and AT&T Bell Labs. He received the University Excellence in Teaching Award at Illinois Institute of Technology in 1999. His research interests include wireless networks and computer systems. Currently, he is an editor of the *Journal of Microprocessors and Microsystems* and *IT Professional*. He is a senior member of the IEEE.

**Ting-Chung Wang** received his M.S. degree in Electrical Engineering from National Yunlin University of Science and Technology, Taiwan, in 2008. He is an engineer with King Yuan Electronics, Taiwan. His current research interests include communication protocols and wireless Internet applications.