

An efficient trigger to improve intra-WiFi handover performance

Roberta Fracchia, Guillaume Vivier

Motorola Labs, Parc les Algorithmes, Saint-Aubin, 91193 Gif-sur-Yvette, France

Abstract Seamless mobility is now a key requirement for wireless communication systems and users are day after day thirstier of high speed connections. For indoor communication, WLAN has become the most common technology. In this paper we introduce a composite trigger for intra-system handover which can take into account not only signal strength but also access point load and expected throughput. Simulation results show that the proposed metric provides gains in term of delays and system overall throughput in various scenarios.

I. INTRODUCTION

Cellular systems are now and from several years fully integrated in the communication landscape. They provide not only wireless connectivity but also seamless mobility: while moving, the users do not experience any interruption in their service. For that purpose, cellular systems introduced a key feature: handover. Seamless mobility has become hence a natural feature of any wireless communication system. In parallel, broadband wireless communication in the home, in office, even in public area (airport, railway station, shopping mall, etc...) has been widely popularized thanks to the WLAN (WiFi). Cheap access points can be easily installed by any individual in its home. Moreover, with the advent of DSL, cable or fiber to the home, customers have access to multiple-play offers, including broadband access to the internet, voice over IP, IPTV... Thus, the WLAN should support more and more delay-constrained services (such as voice) providing an equivalent experience than cellular systems. For that purpose, QoS support has been introduced in WiFi with the amendment "e" of 802.11 as well as several ways to mimic handover between various access points. Moreover, more and more Access Points (AP), mostly offering free of charge services, are becoming available: users could select the less congested or closer AP, thus the AP offering the best connection performance.

Although numerous works have been done to improve the handover process between 802.11 access points (AP), the trigger to handover is often the same: either a connection loss in the basic scheme or simply the signal to interference and noise ratio. In addition, WiFi handover mechanisms often rely on mobile IP: the mobility is then handled at layer 3 which may lead to too high delay for real-time application.

This paper investigates different triggers, aiming at considering expected data rate or traffic load in the different AP. We propose a simple metric that improves the system capacity as well as individual handover delays. Moreover, we compare L2 and L3 handovers in terms of delay. Section 2 presents related work. Section 3 introduces the new metric and describes a practical implementation for evaluating this metric in the terminal side. Simulation results are provided in section 4 for various typical scenarios.

II. Related Works

Handover is a fundamental mechanism to maintain seamless connections during mobility and different works proposing different handover triggers have been published.

Work in [1] proposes to perform a handover when the Received Signal Strength (RSS) on the new cell is higher than that on the old one. Thresholds, hysteresis times and dwell timers are considered.

This mechanism, that lets only to maintain the wireless connection, can be improved, as proposed in [2]: information on speed and mobile user position is used to estimate the time that would be spent in target cell if the handover would be performed. If this time is bigger than a minimum required time, then the handover is triggered. The minimum required time depends on handoff latency, decision latency and on the time necessary to receive, in the target cell, the same amount of data that would have been received in the current cell. However, even if this mechanism takes into account the data rate that could be used, it can not maximize the user performance, since it does not consider the load of each AP.

Conversely, paper [3] utilizes the beacon transmission delay to estimate the available bandwidth towards different APs and decide to hand over, but not the quality of the signal. Similarly, work [4] assumes that every AP broadcast its load, so that each user may decide to hand over considering the load of the new AP. Moreover, neighboring cell lists are broadcasted to facilitate handover. Authors of [5] instead define a 'transition zone' at the border of the cell coverage. When the user enters the transition zone of the current cell, thus leaving its cell, it hands over. When it enters the transition zone of a new cell, it determines if it can achieve a better QoS and, if necessary, it hands over. However this analysis is not periodically performed, thus reducing the gain that can be achieved.

The algorithm we are proposing tries to overcome the limitations of all these works.

III. Handover Trigger

The algorithm we propose combines all the available quality parameters (Signal plus Interference Noise Ratio –SINR-, Packet Error Rate -PER- and network load). In the following, we first describe the proposed trigger and then explain how to derive the metrics used by the trigger.

A. *The proposed Triggers*

We assume the handover is triggered by a “mobility management” process implemented on the mobile device, which controls continuously the performance of the wireless connection, evaluating the link quality and the channel occupancy, constantly monitored by the 802.11 interface, and triggers the handover toward another AP if needed.

The proposed trigger combines the measured physical layer throughput and the network load in order to maximize MAC layer performance. The metric used to activate the trigger is called “residual throughput” and it is defined as: Data Rate * (1 – PER) * (1 – Channel Occupation). By computing Data Rate * (1 – PER) we deduce the instantaneous PHY layer throughput that reflects the link quality. The word “residual” instead means that a part of network resource is already occupied by other users, and then the handover decision is only based on the remaining bandwidth at the user disposal. The handover trigger is based on a comparison between the estimation of the residual throughput on the current cell (namely, Current_residual_throughput) with the one that could be achieved on another cell (namely, Target_residual_throughput). If the ratio between Target_residual_throughput and Current_residual_throughput is bigger than a threshold,

$$\frac{\text{Target_residual_throughput}}{\text{Current_residual_throughput}} > \text{Throughput_margin}$$

then the handover is triggered.

After extensive simulations we decided to set the throughput_margin to 1.1, a value that avoids ping-pong effects and at the same time does not limit the gain that can be achieved with handovers.

B. *Monitoring*

To derive the Data Rate, the Channel Occupation and the PER, used to compute the residual throughput, the User Terminal (UT) performs measurements in

the used cells and on broadcast messages sent by neighbouring APs. Moreover, in order to restrict the number of cells to measure, we assume that the user terminal is supplied with a list of the neighbouring cells, that are the most likely to fulfill the quality requirements at the user terminal location. In this way, measurements are performed only on the cells included in the neighbours list.

To estimate the bandwidth occupancy we consider the WiFi contention-based access mechanism (CSMA/CA): all subscriber stations continuously listen to the channel before competing for channel access. The usage of the channel bandwidth can be approximated as the ratio between the time in which the channel status is busy according to exchanged frames and the considered time interval. The MAC process records the channel busy time, then periodically calculates the channel occupancy percentage and transmits this information to the “mobility management” process.

Moreover, the MAC layer transmits to the “mobility management” process the SINR measured on downlink broadcast packets (Beacon messages) typically sent every 10 to 100ms. The “mobility management” process moreover transforms the measured SINR into the used data rate value, assuming that a SINR based Link Adaptation (LA) algorithm is implemented. A LA mechanism [6] consists in the selection for each packet transmission of the Modulation and Coding Scheme (MCS) most adapted to the instantaneous wireless link quality, and lets to better exploit the channel performance compared to a fixed data rate. Supposing that the best MCS is used for the current SINR, it is possible to know at all times the data rate and the PER (using SINR vs. PER look up tables). The “mobility management”, using this information, evaluates the performance of available networks as described in the previous section.

IV. Simulation Results

In this section we present results obtained with simulation, showing the effectiveness of the proposed handover trigger. We consider different scenarios, with one or more UTs.

We evaluate the performance of the proposed algorithm providing simulation results derived with OPNET, by modifying the simulator distribution to implement the considered handover criteria, including the “mobility management” entity, which continuously controls the performance of wireless connection, and triggers the handover if needed.

We consider the 802.11g amendment [7], which operates in the 2.4GHz licensed band and offers coverage of 70-80 meters with a theoretical throughput of 54Mbps. We implemented also the link adaptation mechanism, varying the data rate from 6 Mbps for a SINR of 5.5 dB to 54 Mbps for a SINR of 28.2 dB.

A. Handover latency

In this section we evaluate the handover latency experienced using the residual throughput trigger. We compare the results with the latency of a basic handover trigger, which does not utilize the neighboring cell lists and is triggered by the lost of the connection.

We considered two different deployment scenarios, with layer-2 and layer-3 handover respectively.

A layer-2 handover is performed when the UT hands over between APs controlled by the same gateway. In this case the IP address doesn't change and, as a consequence, the use of Mobile IP to maintain the connection is not necessary.

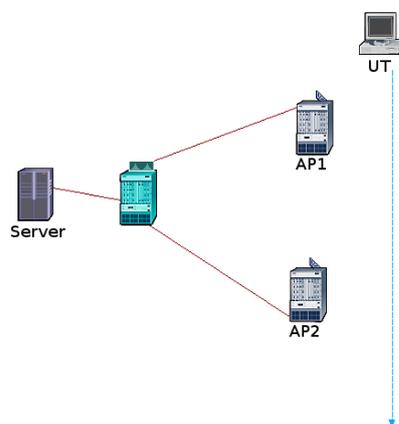


Figure 1. Reference scenario for the evaluation of the trigger without cell congestion

In this context, the handover latency is due to the scanning procedure, the handover decision and the authentication/reassociation phase. Since our simulator does not implement the exchange of signaling messages, the values reported in the figure do not include the authentication/reassociation delay. As reported in several work of measurements in WLANs, that delay usually varies between 5 and 15 ms, depending on the used card. To have a complete overview of the handover latency, these values have to be added to the simulation results.

A layer-3 (plus layer-2) handover is instead performed when the UT hands over between APs controlled by different gateways. In this case, the latency accounts also for the binding updates of the Mobile IP protocol.

The scenario we refer to is represented in Figure 1. Two APs are available and one user terminal, initially connected to AP1 moves away at 5m/s thus entering the coverage of AP2. Different positions of the BSs have been considered: results reported hereafter are averaged over all the performed simulations.

The simulations results are reported in Table 1. We can notice that also in case of layer-2 handover, the latency experienced using the basic trigger is too high to

offer seamless mobility. The delay experienced using the residual throughput trigger instead it is equal to zero, since scanning and measurements are done rapidly and periodically on the cells of the neighbouring lists only and no more at the handover time. We have to remember that the signaling delay has to be added to the values reported in the table.

When Mobile IP has to be used to perform a layer-3 handover, the latency dramatically increases, even if the proposed trigger still offers better performance.

	proposed trigger	basic trigger
Layer2 HO	0	0.432 s
Layer2+Layer3 HO	2.7 s	3.6 s

Table 1: Handover latencies

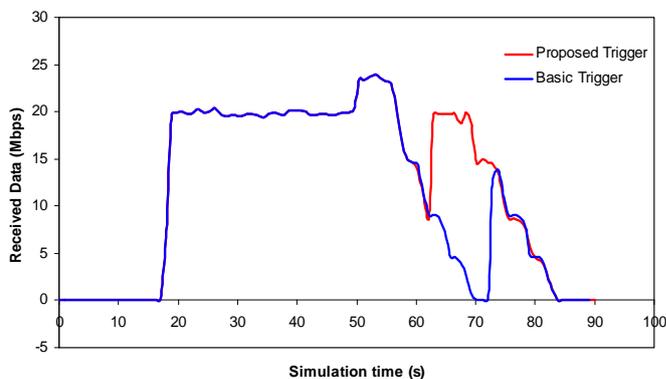


Figure 2. Received data as a function of the simulation time

B. Performance at connectivity's loss

In this section we evaluate the impact of the handover on different applications. We refer to the scenario presented in Figure 1 and we consider long-lived ftp connections generated at the Server node with a rate of 50Mbps.

We report in Figure 2 the amount of received data as a function of the simulation time. We consider again the residual throughput and the basic triggers. As you can see in the figure, the gain achieved with the proposed trigger is not only in terms of latency, but also in terms of achieved throughput: the UT, exploiting information already collected, is able to select the AP that can provide better performance. Without attending the loss of coverage, as soon as the target residual throughput increases the current residual throughput, the handover is performed.

Thus, the handover to another cell allows the use of higher data rates, related to the higher SINR, since an adaptive Modulation and Coding Scheme is adopted.

Finally we consider an interactive application like Voice over IP and we study how the residual throughput trigger impacts on the quality of the communication.

Figure 3 reports the packet delay at the application layer as a function of the simulation time. At the beginning of the communication, which is at 13 s, the UT is in the coverage of AP1 and the delay of the voice packets is about 80 ms. From that moment on, the UT moves in directions of AP2. In the case of the basic trigger, the handover to AP2 is triggered only when the coverage of AP1 is almost lost, that is at 68 s. The poor quality of the communication before the handover and the time necessary to scan all the channels to find the new AP introduce a high delay: you can observe the peak in the delay, that reaches 500 ms. Then the UT continues to move and exits from the coverage of AP2 at 85 s. There are not others APs, so that at 85 s the communication is interrupted. This explains the new arise in the delay between 80 and 85 s.

The behavior of the UT that uses the proposed trigger is the same. The difference in the experienced delay depends only on the handover from AP1 to AP2. Indeed, with the proposed trigger, the handover is performed immediately at 65 s, without degradation in the communication.

This figure clearly shows that the proposed trigger avoids the peaks in the delay due to the interruption of the communication in the classical handover process.

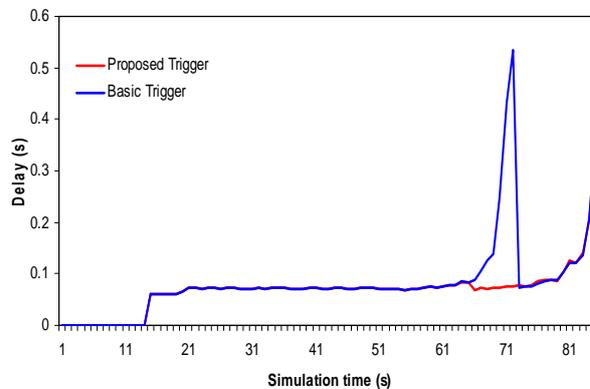


Figure 3. VoIP delay

C. Performance in Congested Scenarios

In this section we show the performance gain that can be achieved performing handovers triggered by the residual throughput in different load conditions.

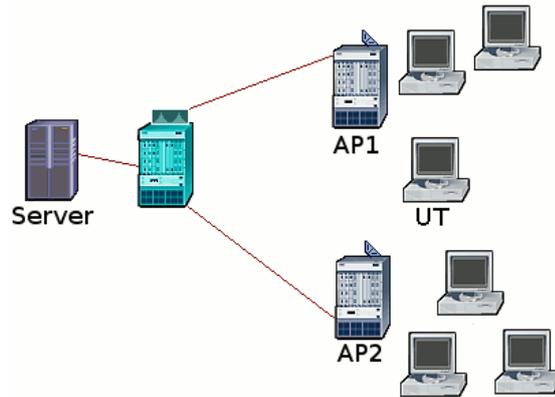


Figure 4. Reference scenario for the evaluation of the trigger with cell congestion

Consider as an example the topology presented in Figure 4. The user terminal UT is under the coverage of both AP1 and AP2 and it is initially connected to AP1. Assume that, due to its placement, the SINR of the radio channel to AP1 perceived by UT is 18.5dB. Then we vary the position of AP2 to achieve SINR in the second cell equal, lower and higher than that in the first cell. Moreover, we assume the other nodes connected to AP1 generate a load of 5 Mbps, while the load of AP2 is varying.

Figure 5 reports the layer 2 throughput achieved by station UT, as a function of the load ratio between the two cells, for different values of the SINR on the target channel (to AP2). The throughput achieved transmitting towards AP1, that is, if no handover is performed, is reported as a reference.

First of all, note that all the handovers result in a throughput gain. The estimation of the residual throughput indeed let the user correctly judge if advantages in performance can be achieved handing over. Moreover, we can gather that when the SINR on the target cell is higher than that one the current cell, the UT can profit of the handover until high values of load of the target cells. Otherwise, if the SINR and thus in the data rate that can be used are lower than that in the current cell, the handover is performed only when the load on the target cell is lower than that on the current cell.

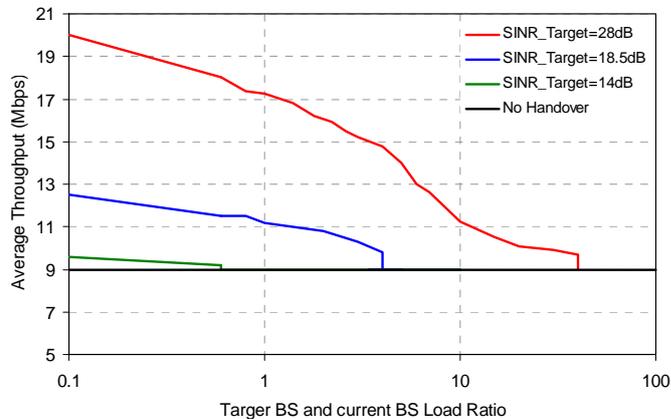


Figure 5. Average throughput of the UT as a function of the load on the target channel for different SINR values

V. Conclusions

In this paper we have simulated the handover between 2 WiFi access points. An efficient but simple metric to trigger the handover has been proposed to take into account user signal quality but also access point load. A practical mean to evaluate the metric was introduced, based on the 802.11 inherent mechanisms. Simulation performance shows that the proposed metric improves handover latency and system throughput, compared to basic handover mechanism.

Although this work concentrates on intra-WLAN handover, the metric can be easily extended to inter-system handover, when for instance WLAN and cellular system are collocated.

Acknowledgments This work has been partially performed in the framework of the IST project IST-4-027756 WINNER II, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues.

References

- [1] G.P. Pollini, Trends in handover design, IEEE Communication Magazine, Mar. 1996
- [2] A. Saleh, A location-aided decision algorithm for handoff across heterogeneous wireless overlay networks, Master Thesis. Virginia Polytechnic Institute and State University, 2004
- [3] S. Vasudevan, K. Papagiannaki, C. Diot, J. Kurose, and D. Towsley, "Facilitating Access Point Selection in IEEE 802.11 Wireless Networks," in ACM Sigcomm IMC, Oct. 2005.
- [4] W. Zhang, J. Jähnert, K. Dolzer, Design and Evaluation of A Handover Decision Strategy for 4th Generation Mobile Networks, IEEE VTC, April 2003, Jeju, Korea.
- [5] S. Balasubramaniam and J. Indulska, Vertical handover supporting pervasive computing in future wireless networks, Computer Communications, Volume 27, Issue 8, May 2004

- [6] Goldsmith A.J., Chua S.-G., "Adaptive coded modulation for fading channels", IEEE Transactions on Communications, Volume 46, Issue 5, May 1998
- [7] IEEE Computer Society, IEEE 802.11 Standard, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications", November 2001