Experimental Study of Mismatching ESS-Subnet Handoffs on IP over IEEE 802.11 WLANs

Li-Hsing Yen*, Hung-Hsin Chang†, Shiao-Li Tsao‡, Chia-Chi Hung‡ and Chien-Chao Tseng‡
*Dept. of Computer Science & Information Engineering
National University of Kaohsiung, Kaohsiung, Taiwan 811, R.O.C.
†Dept. of Information Management, Chin Min Institute of Technology, Toufen, Miaoli, Taiwan 351, R.O.C.
‡Dept. of Computer Science, National Chiao-Tung University, Hsinchu, Taiwan 300, R.O.C.

***Abstract*—To support delay-sensitive applications on IP over wireless LAN (WLAN), both layer-2 handoff (L2H) and layer-3 handoff (L3H) must be conducted efficiently. Prior studies toward a fast L2H/L3H assume simple networking environments where Extended Service Set (ESS) exactly matches subnets. This paper identifies performance issues associated with inter-ESS L2Hs and L3Hs under mismatching ESS-subnet configurations, and analyzes actual performance impact on existing systems. Experimental results show that inter-ESS L2Hs are much time expensive than intra-ESS L2Hs while mismatching ESS-subnet settings lead to either time-consuming L3H detections or redundant L3H executions. We also discuss possible remedies for this problem.**

I. Introduction

An IEEE 802.11 wireless station (WS) residing in a wireless local area network (WLAN) must associate with some access point (AP), after which the WS can then send traffic through the associated AP to a wired infrastructure. If the associated AP is no longer accessible to the WS for some reason, the WS should discover and attempt an association migration with another AP to continue its ongoing sessions. The process of changing a WS's association between two APs is called *layer-2 handoff* (L2H). If the L2H involves a change of network domains (subnets), a *layer-3 handoff* (L3H) is also needed to renew network settings associated with that link.

An Extended Service Set (ESS) consists of one or more interconnected APs that are configured with the same Service Set Identifier (SSID). ESSs are usually configured in accordance with the hierarchy of access networks to ease network management, but SSIDs and subnets are in fact two independent configuration settings (one layer-2 and the other layer-3).

A typical example of mismatching ESS-subnet configuration is a citywide or national wireless infrastructure that utilizes hundreds or thousands of APs. These APs may be located in different IP subnets, yet they share the same SSID for an easy identification by roaming users. In such case, changing subnets does not entail a change of ESSs. Another mismatching ESS-subnet example can be found on campus networks, where a student may be authorized to access several ESSs that are operated by different laboratories located in the same network segment. In this case, changing ESS does not necessitate a change of subnets.

TABLE I
Handoffs under different ESS-subnet settings. Names in italics are mismatching ESS-subnet handoffs.

L2H	L3H	
	Intra-subnet	Inter-subnet
Intra-ESS	Intra-ESS/Intra-Subnet (aE-aS)	*Intra-ESS/Inter-Subnet* (aE-rS)
Inter-ESS	*Inter-ESS/Intra-Subnet* (rE-aS)	Inter-ESS/Inter-Subnet (rE-rS)

Accordingly, four possible types of handoffs can be defined, each corresponding to a unique combination of ESS and subnet settings. Among them, two are *mismatching ESS-subnet handoffs* (Table I). An intra-ESS/inter-subnet (aE-rS) handoff means that the old and new APs are of the same ESS but located in different subnets, while an inter-ESS/intra-subnet (rE-aS) handoff refers to a change of AP associations that are within the same subnet but not the same ESS.

The contribution of this paper is twofold. First, we have investigated empirically the time cost of inter-ESS L2Hs. In particular, we observed that current implementations of IEEE 802.11 network interface cards (NICs) do not allow inter-ESS L2Hs only if no accessible AP can be found in the current ESS for an extended period of time. The time period is long enough to raise a *link-down* event indicating a (possibly temporal) link breakdown. This design makes inter-ESS L2Hs time expensive.

The other contribution of this work is to demonstrate how the interpretation of link-down event as a signal to start L3H affects L3H performance under mismatching ESS-subnet environment. In the literature, only handoffs under matching ESS-subnet settings have been considered [1], [2]. Many L3H proposals and implementations take link-down and related events as a major (and sometimes the only) facility to signal the need for an L3H. However, the handling of these events in practice does not completely meet the semantic requirement of most L3H proposals/implementations. Our case is exactly an instance of this problem, as a link-down event accompanying with a transition of ESS does not necessarily indicate the need for an L3H. Consequently, an unnecessary L3H is conducted as a side effect of a rE-aS handoff while an L3H essential to an aE-rS handoff may not be realized immediately after an L2H. We analyzed empirically how this problem affects

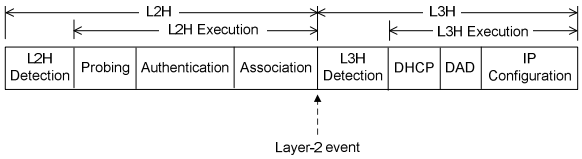

Fig. 1. General L2H/L3H phases

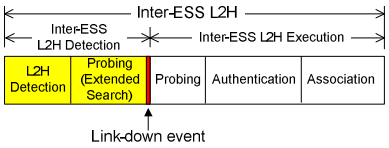

Fig. 2. Handoff stages for inter-ESS L2Hs

L3H latency by performing well-designed experiments with off-the-shelf products from major vendors.

Rest of this paper is organized as follows. Section II briefs technical background of the mismatching problem. Section III presents our observations and numerical results obtained from experiments. In Section IV, we suggest some solutions to this problem. Section V concludes this work.

II. Technical Backgrounds

Figure 1 shows general L2H/L3H phases. An L2H is triggered by L2H detection. L2H detection can be conducted in various ways. A widely-adopted design demands WSs to constantly monitor received signal strength and/or bit error rate from the associated AP. A significant drop of the monitored measure (e.g., below a preset threshold value) is taken as a possible indication of losing the contact with the currently associated AP. Another pure layer-2 approach takes unrecoverable frame transmission errors as an indication. Regardless of how L2H detection is implemented, a critical mission is to tell intermittent bad communication states from persistent ones that must call for an L2H.

Following L2H detection is an L2H execution consisting of probing, authentication, and association phases. The probing phase is usually embodied by an active scan that searches all channels for accessible APs. In active scans, a WS broadcasts Probe Requests on every channel and listens for possible Probe Responses from APs. By default, the WS considers only APs that belong to the current ESS. This is achieved by specifying in Probe Request an SSID (Service Set Identifier) that identifies the designated ESS. Only APs that have been configured with the same SSID should respond when receiving probe requests. If the WS ever receives a response from some AP, indicating a success of the active scan, it can select one AP as the handoff target and proceed to the next L2H phase (authentication). If it further completes the L2H in time, a *link-up* event is signaled to inform upper-layer entities of the change of link.

However, in case that no Probe Responses are received during an active scan, the WS cannot advance to the next phase due to the absence of accessible AP. This eventually leads to a link-down event, which indicates a loss of link connectivity with the current AP. Link-down events are to be captured by software modules beyond layer-2 (L2). One of these modules may instruct L2 to attach to another ESS based on some policy (user preference, for example) to retain network connectivity, effectively conducting an *ESS transition*, namely inter-ESS L2H. Inter-ESS L2Hs are outside the scope of IEEE 802.11. In the mentioned platform, inter-ESS L2Hs are activated only after occurrences of link-down events. Consequently, the time cost of detecting the need for an inter-ESS L2Hs is at least that caused by a failed intra-ESS L2H. Viewing from this perspective, a link-down event divides an inter-ESS L2H into two stages (Fig. 2):

1) Inter-ESS L2H detection, where the need for an inter-ESS L2H is indicated by the failure of an intra-ESS L2H, and
2) Inter-ESS L2H execution, where a new ESS is found and an association with that ESS is made.

Activities in the second stage are identical to those in an ordinary intra-ESS L2H execution. However, how WS acts in the first stage varies in existing implementations. WS may simply keep searching for the designated ESS. It is also common that WS executes an *extended search* which explores potentially accessible APs belonging to other ESSs. One goal of our research is to investigate empirically the impact of ESS transitions on overall L2H latency, and see if different inter-ESS L2H detection designs improve the result.

The other goal of this research is to investigate the impact of mismatching ESS-subnet configuration on L3H latency. The aim of an L3H is to retain network-layer connectivity. In case of IP networks, the center of an L3H execution is to renew network-layer settings by performing Dynamic Host Configuration Protocol (DHCP), possibly Duplicate Address Detection (DAD), and then IP configuration.

L3H detection is the procedure that judges the need for an L3H following an L2H. It is the key to the time gap between the completion of an L2H and the commencement of the succeeding L3H. However, there is no standard way to perform L3H detections. Many existing schemes propose using some L2 event as a trigger to L3Hs. Such event is referred to as an L2 trigger [3]. An L2 event is an ideal L2 trigger if the occurrence of this event always entails the need for an L3H. A link-up event simply indicates an attachment to a new AP. It calls for an L3H only in case of inter-subnet handoffs. Therefore, link-up event alone is not an ideal trigger for L3Hs. On the other hand, a link-down event in practice simply indicates a loss of link connectivity with the current AP. It necessitates an L3H only if the loss of link connectivity comes from a change of access network across different network domains. For instance, a *seamless* L2H (by "seamless" we mean an L2H without temporary link breakdowns) between heterogeneous access networks (one WiFi and the other WiMAX, for example) is currently not possible in reality. A vertical handoff thus always entails a link down event which can be used to trigger an L3H. However, link-down events may also arise for other causes.

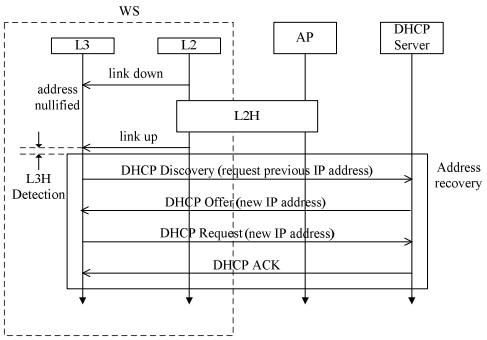

Fig. 3. An L2-triggered L3H in the case of rE-rS handoff

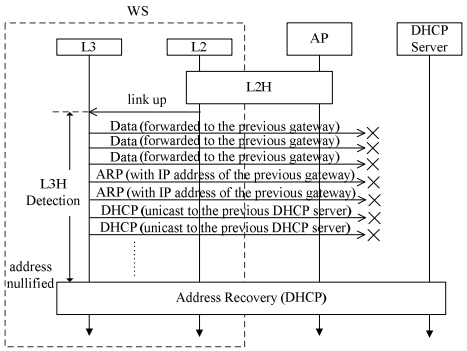

Fig. 4. An L3-initiated L3H in the case of aE-rS handoff

ESS transition is one such cause that is of interest in this paper. Furthermore, it is also possible that an L3H should be performed while no link-down event is signaled. An aE-rS handoff is one such example. Therefore, link-down event is neither an ideal L2 trigger. Unfortunately, link-down or similar event (e.g., Link_Going_Down [4], Link-To-Be-Down [5]) has been misinterpreted as an appropriate (and sometimes the only) L2 trigger. In this paper, we shall demonstrate how L3H latency is affected by such misuse with the example of mismatching ESS-subnet configurations.

III. An Experimental Study

A. Observations

As we are interested in actual performance in real systems, it is crucial to understand how L3H is dealt with in practice. In Windows XP, we observed that the need for an L3H is realized through a two-step process:

1) Firstly, address settings associated with some link get nullified. The nullification of address settings may come from either a link-down event or excessive losses of L3 packets on that link. We call L3Hs that are caused by the former *L2-triggered* L3Hs while those caused by the latter *L3-initiated* L3Hs.
2) Secondly, the layer-3 (L3) module attempts to recover the previous IP address by issuing DHCP Discovery messages with option "Requested IP Address" set to the previous one. If this recovery attempt succeeds, the L3 module assesses that it remains in the same subnet and no L3H is needed. On the other hand, the failure of the address recovery attempt concludes the need to execute an L3H.

In L2-triggered L3Hs, these two steps are not continuous. After a link-down event nullifies address settings, the L3 module cannot proceed with the address recovery attempt due to the lack of an active link. This means that these two steps are separated by an L2H. Fig. 3 shows a timing diagram for L2-triggered L3Hs which typically occur for rE-rS handoffs.

An L3-initiated L3H can be created by an association migration across different subnets without raising a link-down event (e.g., an intra-ESS L2H.) For this type of association migration, address settings do not change after L2H and the L3 module keeps forwarding outgoing packets to the default gateway associated with the previous subnet. However, these packets are not deliverable in the new subnet. Meanwhile, ARP (Address Resolution Protocol) cache entries associated with the previous subnet cannot get refreshed and eventually age out. Subsequent outgoing packets cause a series of L3 error-resolving activities:

- The L3 module issues ARP requests for the IP address associated with the previous gateway.
- The L3 module unicasts DHCP request messages toward the previous DHCP server to renew the lease of the WS's IP address.

The failure of these error-resolving actions eventually clears address settings associated with the previous subnet. The address recovery attempt then follows immediately. Fig. 4 shows a timing diagram for an L3H-initiated L3H caused by an aE-rS handoff.

It is not hard to comprehend that L3H detection in L2-triggered L3Hs should be faster than that in L3-initiated L3Hs. Therefore, the L3H-detection rule mentioned above suits well to networks with matching ESS-subnet settings, where L3H is needed if and only if the WS undergoes an ESS transition, and an ESS transition always entails a link-down event and accompanying L2-triggered L3H. In networks with mismatching ESS-subnet settings, however, the rule may lead to poor performance. In rE-aS handoffs, ESS transitions give rise to link-down events, which nullify address settings and in turn trigger L3H executions. These L3H executions are redundant for rE-aS handoffs. For aE-rS handoffs, L2H can be done without raising a link down event and address settings usually remain intact after L2H. WSs can only resort to L3-initiated L3Hs for the change of subnets, which could be very time-consuming. Such phenomenon has not been studied previously.

B. Experiments

We conducted experiments to measure handoff latencies under various ESS-subnet settings. Two IEEE 802.11g APs were deployed, one operating at channel 1 and the other channel 6. We replaced each AP's external antenna by a cable to direct its signal into a tunable signal attenuator[1],

[1]E-Instrument Tech LTD. Model: EPA-1200

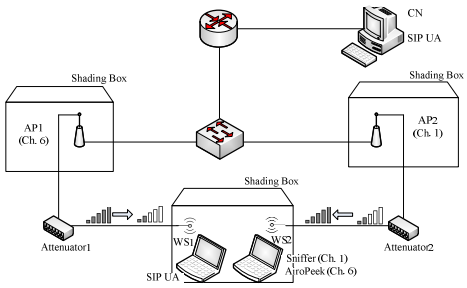

Fig. 5. Experimental setup for intra-subnet handoffs

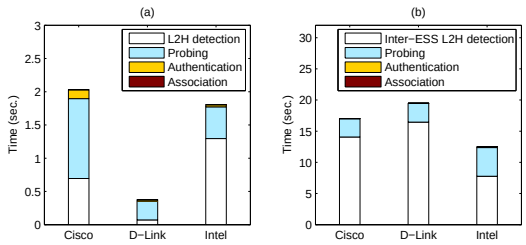

Fig. 6. Delays of (a) intra-ESS L2Hs and (b) inter-ESS L2Hs

TABLE II
HARDWARE EQUIPMENT LIST

Name	Type	Model	Chip
AP1, AP2	AP	ZyXEL P-330W	Realtek (RTL8186)
Cisco	NIC	Cisco Aironet AIR-CB21AG-W-KG	Athores
D-Link	NIC	D-Link DWL-G650	Athores
Intel	NIC	Intel(R)PRO/Wireless LAN 2100 3B Mini PCI	Intel

from which the signal is further fed into a signal-shading box[2]. In the shading box two WSs were placed: WS1 was the one to exercise handoffs while WS2 was an observer. The observer was equipped with two 802.11g interfaces operating in promiscuous mode. It captured all frames exchanged on channels 1 and 6 using Sniffer and AiroPeek, one for each interface. Both WS1 and a corresponding node (CN) located in another subnet were installed a SIP UA (User Agent). This was to establish Voice over IP (VoIP) sessions between them. The UA software had been modified so that it can capture link-down and link-up events from NIC drivers. We also ran Ethereal on WS1 and the CN to capture all outgoing SIP-related packets. The experimental setup for intra-subnet handoffs is shown in Fig. 5. For inter-subnet handoffs, the only difference is that AP1 and AP2 were connected by a router instead of a switch.

We let WS1 associate with AP1 initially. After the association, WS1 used SIP to establish a VoIP session with the CN. The attenuators were then used to simulate a handoff of WS1 from AP1 to AP2. The distance between AP1 and AP2 was assumed 50 m. For this setting, the signal power of AP1 was linearly decreased from 0 dBm to -74 dBm with the decreasing rate set to 1 dB per second. Meanwhile, the signal strength of AP2 was linearly increased from -74 dBm to 0 dBm with an increasing rate of 1 dB/s. Note that the decreasing and increasing rate of signal strength affects only the latency of L2H detection.

WS1 was running Windows XP, in which built-in ZeroConfig was used to enable inter-ESS handoffs. To ensure that the obtained results are common to a variety of products, WS1 was equipped with three types of NICs as listed in Table II.

Our captured traces show that when signal strength degraded to some level, transmission failures occurred to frames at times. Retransmissions of these frames might succeed or fail, and we took the first retransmission failure as the start point of L2H detection. Note that frames transmitted after this point might still be transmitted successfully. For intra-ESS L2Hs, the issue of the first Probe Request was considered the end of L2H detection phase and also the start point of the probing phase. For inter-ESS L2Hs, link-down events further separated the detection and the execution of inter-ESS L2Hs.

Fig. 6 shows measured handoff latency for every NIC in both intra-ESS and inter-ESS L2Hs. Each measure stands for the average of data collected from at least 25 runs.

It can be seen that the delays of inter-ESS L2Hs are generally much higher than those of intra-ESS L2Hs. The delays of authentication and association phases are nearly the same in both types of L2Hs. It is handoff detection and probing that dominates the results. We also confirmed that NICs did not initiate an inter-ESS L2H until the occurrence of a link-down event. It is the timeout setting of link-down event that dominates the length of inter-ESS L2H detection.

We also observed two types of designs in inter-ESS L2H detections. One design is to always perform an intra-ESS probe (SSID-specific probe) and an inter-ESS probe (broadcast probe) simultaneously during a regular probing. This strategy helps early explorations of APs in other ESSs (though this was not necessary for intra-ESS L2Hs). D-Link NIC adopts this design approach. The other design is to conduct an extended search only after a regular probing fails. Intel and Cisco NICs take this design. Intel NIC implements extended search by keeping broadcasting probe requests, and these requests are alternatively destined for the original ESS and any ESS (not specifying a specific SSID). Cisco NIC performs extended search by sending to each AP a Probe Request with SSID set to the original ESS using a unicast frame. Since the request is a unicast, every AP successfully receiving it responds an ACK frame, even if the SSID of the AP does not match the target. It should also be a sort of intra- and inter-ESS probe.

Although all these NICs could exploit extended search results that were acquired before link downs, there is no evidence that any of them actually did it since all these NICs performed another active scan after link downs. As a result, extended searches neither shorten inter-ESS L2H detection time nor save the failure of intra-ESS L2Hs.

For L3H latencies in different ESS-subnet configurations, Table III summarizes our experimental results (aE-aS handoffs do not incur any L3H cost and are therefore not shown here.)

[2]Augleton Tech. Inc., model:581-200

TABLE III
L3H LATENCY (UNIT: SECOND)

Setting	NIC	L3H Detection	L3H Execution	Total
Inter-ESS/Intra-Subnet	Cisco	0.090	3.034	3.124
	D-Link	0.086	3.047	3.133
	Intel	0.141	3.332	3.473
Intra-ESS/Inter-Subnet	Cisco	8.928	2.768	11.696
	D-Link	9.196	3.066	12.262
	Intel	8.167	3.008	11.175
Inter-ESS/Inter-Subnet	Cisco	0.087	3.176	3.263
	D-Link	0.083	3.141	3.224
	Intel	0.056	3.754	3.810

Each measure stands for the average of data collected from at least 25 runs. In the table, L3H detection latency counts from the completion of L2H to the start of the standard L3H execution (the first broadcast of DHCP Discovery messages).

In case of rE-aS handoffs, all NICs signaled link-down events, which triggered redundant L3Hs. Redundant L3Hs cost 3.0 to 3.3 seconds according to our measurements. This amount is significant for real-time applications such as VoIP.

For aE-rS handoffs, which were L3-initiated L3Hs, the time for L3H detection dominated overall L3H latency. The observed L3H detection latency ranged from 8.2 to 9.2 seconds. In contrast, for rE-rS handoffs, which were L2-triggered L3Hs, the observed L3H detection time was much shorter (generally less than 0.1 second).

IV. Discussions

We summarize the experimental results with the following four points:

- Inter-ESS L2Hs are much time expensive than intra-ESS L2Hs. This is because inter-ESS L2Hs are carried out only after the failure of intra-ESS L2Hs.
- L2-triggered L3Hs had shorter L3H detection time than L3-initiated L3Hs. Therefore, link-down events as an L2 trigger are beneficial to L3H detection with matching ESS-subnet settings.
- One the other hand, link-down events as an L2 trigger under mismatching ESS-network settings gave rise to redundant L3Hs (about three seconds).
- For aE-rS handoffs, the use of link-down events as an L2 trigger did not help reduce L3H detection time. The resultant L3-initiated L3Hs had an L3H detection latency ranging from eight to nine seconds.

About the first point, an attempt to reduce inter-ESS L2H time cost by minimizing the link-down timeout value is not feasible as this may increase the probability of unnecessary ESS transitions. Alternatively, we may modify the behavior of L2 entity to enable ESS transitions as soon as the information of APs in another ESS is acquired without raising link-down events. However, this involves the acquisition of information from higher layer such as security context and user preference in ESS.

One solution to the last two issues is to totally preclude mismatching ESS-network deployments. However, this policy may not be feasible under some circumstances. More importantly, link-down events as an L2 trigger may occur for other reasons not calling for an L3H. After all, it simply indicates a loss of link connectivity with the current AP, not necessarily a need for L3Hs. Therefore, a more thorough solution is to devise an L2 trigger that exactly captures AP-subnet relationship even under mismatching ESS-network environments. The L2 trigger could be, for instance, a link-up event accompanied with network information [6]. To this end, NIC drivers should be able to attain AP-subnet mapping information during the execution of an L2H. The mapping, however, is a kind of cross-layer topological information [7] beyond L2 and is not yet available in today's wireless infrastructure. Tseng *et al.* [7] suggest maintaining a dedicated server for the association information between APs and Mobile IP Mobility Agents, as well as for the location and neighborhood information of APs. We could easily extend the contents of such store to include additional AP-subnet mapping information. The work in [8] lets each WS locally cache AP-subnet mapping information for each AP it has visited so that it could acquire subnet information immediately via the cache upon revisiting the same AP. Yet another way is to include the subnet address of AP in Beacon or Probe Response frame. Whichever approach is taken, an ideal (but not yet realized) L2 trigger could eliminate lengthy L3H detection as well as redundant L3Hs.

V. Conclusions

We have identified performance issues associated with inter-ESS L2Hs and L3Hs under mismatching ESS-subnet configurations, and analyzed actual time performance of existing systems. The experimental results have revealed that 1) inter-ESS L2Hs are much time expensive than intra-ESS L2Hs and 2) the mismatching problem gives rise to either time-consuming L3H detections or redundant L3H executions. Possible solutions to the mismatching problem have been discussed.

References

[1] A. Mishra, M. Shin, and W. 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, Apr. 2003.

[2] V. Bychkovsky, B. Hull, A. Miu, H. Balakrishnan, and S. Madden, "A measurement study of vehicular Internet access using In Situ Wi-Fi networks," in *Proc. MobiCom'06*, Sep. 2006, pp. 50–61.

[3] K. El Malki *et al.*, "Low latency handoffs in Mobile IPv4," IETF, Tech. Rep. RFC 4881, Jun. 2007.

[4] S. Tran-Trong, S. Tursunova, and Y.-T. Kim, "Enhanced vertical handover in Mobile IPv6 with Media Independent Handover services and advance Duplicate Address Detection," in *KNOM Conference*, 2008.

[5] S. M. Yoon, S. J. Yu, and J. S. Song, "Cross-layer fast and seamless handoff scheme for 3GPP-WLAN interworking," in *Innovative Algorithms and Techniques in Automation, Industrial Electronics and Telecommunications*, T. Sobh *et al.*, Eds. Springer, 2007, pp. 437–442.

[6] S. Krishnan *et al.*, "Link-layer event notifications for detecting network attachments," IETF, Tech. Rep. RFC 4957, Aug. 2007.

[7] C.-C. Tseng, L.-H. Yen, H.-H. Chang, and K.-C. Hsu, "Topology-aided cross-layer fast handoff designs for IEEE 802.11/Mobile IP environments," *IEEE Commun. Magazine*, vol. 43, no. 12, pp. 156–163, Dec. 2005.

[8] A. G. Forte, S. Shin, and H. Schulzrinne, "Improving layer 3 handoff delay in IEEE 802.11 wireless networks," in *The 2nd International Wireless Internet Conference*, Aug. 2006.