Fast Authentication in Multi-Hop Infrastructure-based Communication

By: Kevin Lee, Jing Deng, Raghuram Sudhaakar


Made available courtesy of Institute of Electrical and Electronic Engineers (IEEE): http://www.dx.doi.org/10.1109/ICC.2014.6883395

***© IEEE. Reprinted with permission. No further reproduction is authorized without written permission from IEEE. This version of the document is not the version of record. Figures and/or pictures may be missing from this format of the document. ***

Abstract:

Multi-hop infrastructure-based communication is expected to play a vital role in supporting high data-rate multimedia access to mobile devices. The advantages are significant in highly mobile scenarios such as intra-vehicular networks. However, mobile nodes in these networks suffer from long authentication delays, which adversely affect the goodput. In this work, we propose two techniques to shorten the initial authentication delay without compromising the authentication process and overall security. One of the techniques, called fast authentication, admits data traffic temporarily through the network to the gateway and the immediate parent node of the joining node presents network-side authentication. The other technique, called prefetch-assisted authentication, allows the authenticated wireless nodes to prefetch and store the authentication vectors of the potential mobile clients. We investigate several unique features of our proposed schemes and find their performance to be suitable for infrastructure-based multi-hop wireless communications.

Keywords: Authentication | Communication system security | Delays | Mobile communication | Prefetching | Wireless communication

Article:

I. INTRODUCTION

The growing consumer demand for mobile Internet access has car manufacturers giving serious thought to incorporating access capabilities in vehicles. Many manufacturers are jointly working with national government agencies to develop solutions aimed at providing high data-rate Internet services to cars. One of the outcomes is a novel type of wireless access called Wireless Access for Vehicular Environment (WAVE) dedicated to vehicle-to-vehicle and vehicle-to-roadside communications. While the major objective has clearly been to improve the overall safety of vehicular traffic, promising traffic management solutions and on-board entertainment
applications are also expected by the different bodies (C2CCC, VII, CALM) and projects (VICS4, CarTALK 2000, NOW5, CarNet, FleetNet) involved in this field. When equipped with WAVE communication devices, cars and roadside units (RSUs) form a highly dynamic network called a Vehicular Ad Hoc Network (VANET), a special kind of Mobile Ad Hoc Networks (MANETs). In VANETs, vehicles communicate with one another through wireless infrastructures to the Internet using a multihop-to-infrastructure routing protocol.

The multihop-to-infrastructure routing protocol is further motivated by the trend that cellular service providers have changed from a fixed monthly fee to a tiered or per bit pricing structure for data usage. At the same time, these service providers have started to roll out their own WiFi networks. As free (or significantly cheaper) WiFi networks become predictably more accessible from vehicles, users will have a strong economic incentive to opportunistically offload data traffic from 3G and 4G links to WiFi links.

Irrespective of the service provider and the mode of service, a critical problem in multi-hop networks, such as VANETs, is security. More specifically, the combination of multiple hops to a trusted device (such as a provider-installed access point) and mobility accentuates the problem of key exchange and distribution. Almost all existing authentication protocols have been designed for situations in which the client device directly connects to a trusted device (e.g., an access point). When applied to the multihop scenario, the duration of the authentication process increases significantly.

EAP-AKA (Extensible Authentication Protocol using Authentication and Key Agreement) is a popular mechanism used in mobile networks [1], [2] and is also extensively used to secure WiFi-based networks. It has established itself as a de-facto standard for authentication and most of EAP-AKA’s security architecture has been adopted in providing security for 4G Long-term Evolution (LTE) networks. Despite the wide acceptance, the standard results in tremendous network overhead and delay. In multihop wireless networks, it is impractical to use the centralized approach where a joining node is authenticated at the backend by an Authentication Center (AuC) or Home Environment (HE) and the response is sent toward the node through relayed nodes. The long delay in node authentication reduces the time in which meaningful data can be sent. The mobility of nodes only worsens it.

This work focuses on methods to lower the authentication delay in multihop infrastructure-based wireless networks. We propose two techniques: one called Fast Authentication, in which data traffic is temporarily admitted through the network to the gateway and the immediate parent node of the joining node presents network-side authentication; the other called Prefetch-Assisted Authentication, in which Authentication Vectors (AVs) of joining nodes are prefetched to the potential parent nodes in order to achieve authentication without long delays caused by wireless communications. Overall, our approach is to pipeline data traffic in the slow wireless links and to pre-compute a set of AVs to ensure that data stream can start to flow as quickly as possible, while maintaining the system’s security.
Our paper is organized as follows: Section II introduces RPL. System model and problem statement are presented in Section III. In Section IV, we detail our proposed techniques. Theoretical analysis is given in Section V, followed by performance evaluation in Section VI. We conclude the work in Section VII.

II. BACKGROUND

We use Routing Protocol for Low power and Lossy Networks (RPL) [3] as the basis for this study. It was originally designed to meet specific requirements in Low power and Lossy Networks (LLNs), such as sensor networks. RPL has been extended for VANET routing as well [4].

RPL is a distance vector routing protocol that builds a Destination Oriented Directed Acyclic Graph (DODAG) using mechanisms that support local/global repair while limiting control traffic. One or more RPL instances can be built over the physical mesh using a set of metrics and constraints. Each RPL instance is represented as a DODAG anchored at the root, called the LBR (LLN Border Router). An RPL instance may provide routes to certain destination prefixes, reachable via the DODAG roots or alternate paths within the DODAG. These roots may operate independently, or may coordinate over a network that is not necessarily as constrained as an LLN.

Three different modes have been designed to accommodate security protection. In the first, called unsecured, RPL control messages are sent without any additional security mechanisms. In the second, called pre-installed, nodes joining an RPL instance have pre-installed keys that enable them to process and generate secured RPL messages. The third mode is called authenticated. In authenticated mode, nodes have pre-installed keys as in pre-installed mode, but the pre-installed key may only be used to join an RPL instance as a leaf. Joining an authenticated RPL instance as a router requires obtaining a key from an authentication authority. The process by which this key is obtained is out of scope for RPL specification. Currently, there has not yet been any authentication or key exchange protocol proposed for RPL.

A published work similar to our approach is [5], which only focused on single-hop wireless networks. It tried to balance the processing loads among the RADIUS server and APs. Instead, we focus on multi-hop infrastructure-based wireless networks.

A more general survey of key exchange mechanism indicates some previous work in this area. An incentive mechanism was investigated in [6]. A wireless LAN service integration architecture based on current wireless LAN hot spots was proposed by Shi et al. [7]. Yang et al. proposed a two-party anonymous authentication for wireless networks; a temporary anonymous certificate key technique was used in [8]. Tsai et al. investigated security weaknesses of roaming [9]. A user authentication and key exchange technique using bilinear pairing was proposed by Wu and Tseng [10]. Fast handover and authentication problem was investigated in [11]. Li et al. proposed techniques to allow WLAN users to access cellular networks [12].
Different to these state-of-the-art works, our approaches are able to shorten the initial authentication delay with the help from the immediate parent node as well as the root node in the tree of wireless nodes. As such, security is not sacrificed while authentication delay can be significantly lowered, at the cost of some communication overheads.

III. SYSTEM MODEL AND PROBLEM STATEMENT

An RPL client is likely to be connected to the infrastructure network provided by Internet service providers through GSM/UMTS security means. Security mechanism for GSM/UMTS has evolved from EAP-SIM, EAP-AKA, to 4G LTE EAP-AKA. Despite the evolution, the underlying security protocol from one version to another is similar. Figure 1 is a summary of LTE 4G EAP-AKA exchange:

![Figure 1. EAP-AKA authentication and key agreement. “Auth.” represents authentication. Five rounds of transmissions including wireless communications are needed in order to finish the EAP-AKA exchange.](image)

At first, an identity request/response message pair is usually exchanged. The peer’s identity response includes either the user’s International Mobile Subscriber Identity (IMSI), or a temporary identity (pseudonym) if identity privacy is in effect. After obtaining the subscriber identity, the VLR (Visitor Location Register)/SGSN (Serving GPRS Support Node) sends the authentication data request to HE (Home Environment)/AuC (Authentication Center) for AVs. Upon receipt of a request from the VLR/SGSN, the HE/AuC sends an ordered array of n AVs (the equivalent of a GSM “triplet”) to the VLR/SGSN specifically for the requesting mobile...
equipment. The AVs are ordered based on sequence number. Each AV consists of the following components: a random number RAND, an expected response XRES, a cipher key CK, an integrity key IK and an authentication token AUTN. Each AV is good for one authentication and key agreement between the VLR/SGSN and the UMTS Subscriber Identity Module (USIM).

When the VLR/SGSN initiates an authentication and key agreement, it selects the next AV from the ordered array and sends the parameters RAND and AUTN to the user. AVs in a particular node are used on a first-in / first-out basis. The USIM checks whether AUTN can be accepted and, if so, produces a response RES which is sent back to the VLR/SGSN. The USIM also computes CK and IK. The VLR/SGSN compares the received RES with XRES. If they match, the VLR/SGSN considers the authentication and key agreement exchange to be successfully completed.

The established keys CK and IK will then be transferred by the USIM and the VLR/SGSN to the entities which perform ciphering and integrity functions.

\[ \text{Figure 2. In a wireless network with root as the WiFi AP. Multiple hops of wireless communications are needed. Node } \text{i} \text{ tries to join the network at level } L = 3. \text{ The parent node } j \text{ is at level 2.} \]

If a typical 4G LTE EAP-AKA exchange is employed in the tree-based multihop-to-infrastructure networks, the authentication delay is proportional to the level of the tree. Consider the tree in Figure 2 and a node i joins at level \( L = 3 \). For each EAP-Response/Identity that arrives at the root, the root will have to send an AV (Challenge) to the joining node. The joining node i will then have to send an RES (Response) before it can be authenticated and its data traffic can pass through the network to the Internet. Assume that the delay from one hop to another is one unit, the total delay experienced by node i for authentication is \( 5L \) (EAP-Request, EAP-Response/Identity, Challenge, Response, and Success). Such a security exchange is obviously not scalable to network size as it grows with the size/depth of the network. It does not work for multi-hop to infrastructure network, leading to the so-called Wireless Multihop Authentication problem:
Problem Statement: Wireless Multihop Authentication In wireless multihop networks, authentications using a typical 4G LTE EAP-AKA exchange entail long delays because of the multiple rounds of information exchanges. What technique(s) can be used to reduce such delay while maintaining the same security protection level?

We propose our scheme to reduce the delay in detail in Section IV. The goal is to improve authentication efficiency by reducing the number of round trips before data transmission is allowed.

IV. PROPOSED SOLUTION

We propose two methods of fast authentication and prefetchassisted authentication aiming at reducing authentication delay by pushing authentication procedure to trusted nodes and prefetched AVs, respectively.

A. Fast Authentication

In Fast Authentication, traffic is temporarily admitted through the network to the gateway. Nodes that are already in the network will help HE/HLR authenticate the joining node by sending an EAP-Request with a RAND and a MAC (Message Authentication Code) in the form of \( \{\text{RAND}||\text{MAC}\}\ K\ \text{GRP} \), where \( K\ \text{GRP} \) is the group key of the network. The MAC can also be signed by the key shared between the joining node and HE/HLR and sent along on the downlink traffic. With the MAC, the joining node will be able to authenticate the network that it is trying to join. A replay attack is not possible because RAND is chosen randomly. It then uses the shared key with HE/HLR to generate the Response and compute IK and CK. The joining node will encrypt its data traffic and identity along with MAC by the key derived from CK and IK. The parent will send the traffic along with RAND it generated as a challenge to the child and the child’s identity. HE/HLR will be able to obtain the key shared with the child based on the RAND, derive CK and IK, verify, and decrypt the data traffic. The number of messages for authentication is 2 (EAP Request/Challenge and Response).
Fast Authentication pushes part of the authentication procedure from HE/HLR to the nodes in the network. Instead of receiving a Challenge all the way from the HE/HLR, it is coming from its parent (e.g., node i will be authenticated by node j in Figure 2). The capability of generating the RAND as part of the authentication procedure is given to nodes which are authenticated level by level into the network. Since authentication of the node is partly taken care of by nodes that are already authenticated into the network and traffic is temporarily admitted along with authentication Response, authentication delay is cut to a constant.

B. PreFetch-Assisted Authentication

In Prefetch-Assisted Authentication, the root of the tree prefetches the AVs of the potential clients from AuC. The list of such potential clients can be obtained from roots of the neighboring trees as these mobile nodes are the likely ones that may join the tree in the near future. This technique works well for mobile nodes moving from one tree to another with directions. For the newly powered-up nodes, they can use the Fast Authentication mechanism described above. In fact, users of these mobile devices usually expect long authentication/connection delays. Note that a potential client only needs one AV item instead of n items. This is because any fast authentication would be a temporary one and a normal authentication should be performed later on as EAP-AKA regulates.
Figure 4. Prefetch-Assisted Authentication. The authenticated wireless nodes store the AVs of some potential mobile clients before these clients initiate an authentication exchange. The joining client will authenticate with its parent node with the help of the stored AV.

Once the AVs are prefetched, they will be distributed to all the authenticated nodes in the tree. As an illustration, node i in Figure 2 moves into the range of a tree and becomes a new child of node j, which has been authenticated and carries AVs of some potential clients include node i. Node j can perform mutual authenticate with node i using the prefetched AV value of node i. Figure 4 illustrates Prefetch-Assisted Authentication.

An SQN value is used to ensure AV sequence. Mobile nodes should reject AV values with SQN values too big or too small from the current one that has been recorded since the last authentication. The AV prefetching technique could theoretically disrupt such a sequence. The acceptance window of different SQN’s should be related to the prefetching frequency as well as the size of the AV list for each potential mobile client.

V. SYSTEM ANALYSIS

We present our system analysis of the proposed techniques in this section.

A. Security Analysis
The most important requirement for networks such as VANETs is to prevent unauthorized access. It should prevent malicious users from affecting the performance of the network. In the following, we discuss the performance of our fast authentication and prefetch-assisted authentication scheme in this regard. Note that, we do not focus on users who gain authorized access to the network with the intent of performing malicious activities.

First we discuss the effects of compromised nodes on the network access. For instance, assume that node j has been compromised. It cannot fabricate any AV value to authenticate another collaborating compromised node i that should not be authenticated since the VLR/SGSN will not match the RES(i) and XRES(i) values and will thus reject the request.

Note that a compromised node j can stop a well-behaving child node i from joining the network by modifying the AV and causing checks at the VLR/SGSN to fail. However, the compromised node might not have been authenticated at all, as discussed below.

Secondly, since the proposed methods use that same message exchange techniques as the EAP-AKA, we can guarantee that a malicious node attempting to gain access will fail as already proven in the EAP-AKA mechanism.

B. Delay Analysis

We analyze the authentication delay of our Fast Authentication scheme (see Section IV-A) in the following. Based on the definition of authentication delay, a mobile node at any tree level will experience a fixed authentication delay of $2\gamma$, where $\gamma$ is the transmission of a one-hop message.

As comparison, a regular authentication/joining technique will have an authentication delay of $2L \cdot \gamma$, where L is the tree level at which the mobile node joins (direct child of the root has a level value of 1, etc.) and $L = 1, 2, or 3$, where we have assumed a maximum tree level of 3.

In order to derive such a function, we need to look at the different communication ranges of the WiFi AP node and all other mobile devices. It is foreseeable that WiFi AP node may have a longer transmission range than regular mobile nodes, which try to save energy as well as not to interfere with too many neighbors. Assume that WiFi AP has a wireless communication range of RAP and mobile nodes have a wireless communication range of R. $\text{RAP} = aR$ where $a \geq 1$. 
Highway model: We first study the network scenario of a highway with an AP installed on the road side similar to an RSU [13]. This is basically a one-dimensional model. Assume that the number of cars on each unit distance is uniform and nodes are always connecting to AP with shortest-possible hops. The segment of highway that is within the direct transmission range of AP is 2aR and each additional hop covers 2R distance (from left or right side of AP, see Figure 5).

Therefore, the probability function of nodes at different levels is simply

\[ P_{r^{(1D)}}(L) = \begin{cases} \frac{aR}{3R+aR} & L=0 \\ \frac{R}{3R+aR} & L=1, 2, 3 \end{cases} \]  

(1)

The probability is basically the chance of throwing a dart on the 1-D band and landing it at different distances from the center of the band.

Based on an authentication delay of 2L\gamma when a joining node’s new parent is at level L, we have the expected authentication delay for the new node as

\[ T^{(1D)} = \sum_{L=0}^{3} P_{r^{(1D)}}(L) \cdot 5L\gamma \]
Figure 6. 2-Dimensional model in network scenario. AP is installed in a busy downtown district as mobile devices from all different locations access Internet by multihop connections through AP. AP’s transmission range is $\text{RAP} = aR$ and each mobile device’s forwarding range as $R$.

Two-Dimensional model: APs can be installed in busy streets instead of beside highways. A simple model for such scenarios is the 2-Dimensional model that can capture mobile devices trying to connect to Internet through the AP from different locations (see Figure 6).

Assuming a simple even-distribution of mobile devices on the 2-D region and, without loss of generality, assuming one device on each unit area. The number of mobile nodes in level $L$ of the network rooted by AP is

$$P_r^{(2D)}(L) = \begin{cases} \frac{\pi (aR)^2}{\pi (3R+aR)^2} & L=0 \\ \frac{\pi ((L-1)R+aR)^2}{\pi (3R+aR)^2} - \frac{\pi (LR+aR)^2}{\pi (3R+aR)^2} & L=1, 2, 3 \end{cases}$$

Comparing (1) and (2), we notice that these two equations are similar except the difference in the power of $2$. Define a system parameter called density dimension, $\lambda$. Density dimension basically captures the behavior of number of nodes at different hop distance toward AP. In 1-Dimensional network such as highways, $\lambda = 1$; in 2-Dimensional network such as downtown district, $\lambda = 2$. In other more realistic networks, $1 \leq \lambda \leq 2$. We expect $2 \leq \lambda \leq 3$ in 3-Dimensional networks such as oceanic sensor networks or high-rise building ad hoc networks.

Summarizing (1) and (2), we get

$$P_r^{(\lambda)}(L) = \begin{cases} \frac{(aR)^\lambda}{(3R+aR)^\lambda} & L=0 \\ \frac{(LR+aR)^\lambda - ((L-1)R+aR)^\lambda}{(3R+aR)^\lambda} & L=1, 2, 3 \end{cases}$$
Based on the assumption of a uniform parent selection, the average authentication delay of a joining node is then (as compared to the 2 authentication delay of our proposed scheme)

\[
T^{(\lambda)} = \sum_{L=0}^{3} P r^{(\lambda)} (L) \cdot 5 L \gamma
\]

We define an efficiency, \( \eta \), as the ratio of delays without and with our proposed scheme,

\[
\eta = \frac{T^{(\lambda)}}{2 \gamma} = \sum_{L=1}^{3} \frac{(L + a)^{\lambda} - ((L - 1) + a)^{\lambda}}{(3 + a)^{\lambda}} 2.5L . \quad (4)
\]

VI. PERFORMANCE EVALUATION

In this section, we present our results in performance evaluation. We first calculate numerical results based on (4) and show the difference in delays with or without our proposed scheme. Then we use a set of street traffic data to compare the delay difference.

**Figure 7** Comparison of delays without and with our proposed authentication technique. We used Fast Authentication (Section IV-A) in the comparison.

In Figure 7, we show the ratio of delays between systems without and with our proposed authentication technique under different network scenarios. It can be seen that, as the difference
in transmission ranges of AP and regular mobile nodes increases, the ratio in delay is smaller. This is because of the higher chance of the joining nodes finding the root (AP) as the new parent, leading to lower authentication delay. The delay ratio is generally smaller for smaller (e.g., 1-D network scenario) with higher chances of finding new parents further away. Overall, the saving of authentication delay with the use of our proposed scheme is in the range of 2- to 5-fold.

In order to compare authentication delay difference in a realistic environment. We use the data collected by the Lab for Software Technology at ETH [14] (file named ct-unterstrass-1day filt.0 adj.mov). The movements were recorded for 300 seconds in the Unterstrass region in Zurich, Switzerland. In the period of time, there are 1907 vehicles involved in the region of roughly $3010 \times 5010$ m$^2$ (see Figure 8). Several fictional APs are superimposed in the region at fixed location. We use the movement record of these vehicles to compute a dynamic neighbor matrix and find the level of point of access by different joining vehicles close to the APs. These numbers are then fed into (4) to compute the difference in delays. The results are shown in Figure 9. The saving in delay can been as even higher. The discrepancy between these results and our analytical results in Figure 7 might have been caused by the clustered arrival of moving vehicles. The actual density dimension, $\lambda$, can be estimated at about 3, depending on mobile node transmission range, R. This is not surprising and is in a similar fashion as path loss component of wireless signals $\tau > 2$ in urban settings.
**Figure 8** Map of the monitored region, Unterstrass, Zurich, Switzerland. Virtual APs are superimposed on the map to compute hop counts toward the APs.

**Figure 9** Efficiency of our scheme based on dataset from Unterstrass, Zurich, Switzerland [14]. The efficiency can be seen as even higher. Note the y-axis does not start at zero.

VII. CONCLUDING REMARKS

With the use of WiFi extensions and the so-called multihop-to-infrastructure communications becoming more and more popular, authentication must be performed efficiently and with short delays. Instead, the current EAP-AKA authentication exchange requires multiple rounds of wireless communications on these multihop networks and is unscalable. In this paper, we have presented two techniques to cut down the authentication delay and allow data traffic to be injected in the network with protections. One of the techniques, called fast authentication, admits data traffic temporarily through the network to the gateway and the immediate parent node of the joining node presents network-side authentication. The other technique, called prefetch-assisted authentication, allows the authenticated wireless nodes to prefetch and store the AVs of the potential mobile clients. This is to facilitate authentication between the parent node and the child node without long delays caused by multihop wireless communications.

We have presented system analysis of the proposed techniques and find them suitable for infrastructure-based multihop communications. Realistic traffic flows have been used to evaluate
our scheme. It has been observed that, with our proposed schemes, a significant saving in delay can be achieved.

REFERENCES


