An Efficient and Secure Handover Protocol for IEEE 802.16m Networks

Heba K. Aslan
Informatics Dept., Electronics Research Institute
Cairo, Egypt

Abstract
Mobile WiMAX (Worldwide Interoperability Microwave Access) requires the re-authentication of mobile stations as they change from one base station to another. IEEE 802.16e uses the Extensible Authentication Protocol (EAP) for authentication and key management. This requires about 1000 ms, therefore, it could not support mobile WiMAX applications such as video conference. In the present paper, we propose a protocol that aims to overcome the inter-domain handover problem. The proposed protocol uses the EAP protocol for authentication and key distribution. The proposed protocol is based on the use of hash functions and the Diffie-Hellman protocol to distribute the keys and to avoid the domino effect. The proposed protocol is analyzed using the BAN logic to ensure that it achieves the goals of authentication and key distribution. Furthermore, the proposed protocol is compared with other handover protocols. The comparison shows that proposed protocol outperforms the other protocols.

Keywords: WiMAX Security, Authentication, Handover, 802.16m.

1. Introduction
Mobile WiMAX (Worldwide Interoperability Microwave Access) is a wireless networking system based on the IEEE 802.16e standard [1]. IEEE 802.16e aims to amend 802.16-2004 [2] and to provide mobility. Mobility means that one Mobile Station (MS) can change from one Base Station to another. This requires the re-authentication of the MS. Recently, IEEE issued a new version, IEEE 802.16m [3], as an advanced air interface to meet requirements of the fourth generation (4G) systems. Before accessing the network, mutual authentication between MS and the network must be performed. IEEE 802.16m uses Extensible Authentication Protocol (EAP) for authentication and key management as well as RSA-based authentication protocol. Due to its flexibility, EAP has become the de facto authentication method for 802.16m [4 and 5]. Full EAP authentication latency requires about 1000 ms [6]. Therefore, it could not support mobile WiMAX applications such as video conference or streaming data. It has to be noted that according to latency guidelines identified by WiMAX forum, the latency should not exceed 50 ms in video conferencing or 100 ms in streaming data [7]. Therefore, executing a complete re-authentication in case of handover from one BS to another is not an ideal solution.

Solutions of the handover problem could be classified into two categories: pre-authentication and re-authentication. In pre-authentication handover [8, 9 and 10], MSs can pre-authenticate to the most likely BSs in order to reduce the handover authentication time. Although this solution solves the latency problem, it requires a large computation and communication overheads. In literature, many solutions have been proposed to solve the re-authentication problem. These solutions are divided into two categories [11]: public key-based schemes and symmetric key-based schemes. The public key-based schemes suffer from a high computation cost. On the other hand, symmetric key-based schemes are divided into three categories: key hierarchy, credential tickets and Security Context Transfer (SCT) mechanisms. Key hierarchy scheme is based on using key hierarchy designed especially for handover purposes. This solution reduces the handover latency but not to the extent that could be suitable for real time applications. On the other hand, some credential tickets schemes are based on incorporating a third party to securely distribute keys (such as Kerberos protocol). Other credential tickets schemes are based on using public key cryptography which results in a high computation cost. In SCT mechanisms, the security context (such as negotiated key) is transferred to a target BS using Context Transfer Protocol (CXTP). These schemes suffer from the domino effect problem in which the compromise of one BS will result in compromising other BS’s that share the same security context. All the previous schemes solve the problem for intra-domain handover, where all BS’s are in the same area. For inter-domain handover, the MS must perform the full EAP authentication protocol which leads to an impractical latency for mobile WiMAX applications.

In the present paper, a protocol that aims to overcome the inter-domain handover problem is proposed. The proposed protocol uses the Extensible Authentication Protocol (EAP) for authentication and key distribution. It is based
on the use of hash functions and the Diffie-Hellman protocol to distribute the keys between mobile stations and base stations. In order to avoid the domino effect, the Diffie-Hellman components are distributed instead of the authentication key itself. The proposed protocol is analyzed using the BAN logic to ensure that it achieves the goals of authentication and key distribution. The analysis shows that the proposed protocol achieves its goals without bugs or redundancies. Furthermore, the proposed protocol is compared with other handover protocols. The comparison shows that the proposed protocol outperforms the other protocols. The paper is organized as follows: in Section 2, a survey of related work is detailed. Then, a description of the proposed protocol is given in Section 3. Next, logical analysis of the proposed protocol is illustrated in Section 4. Then, a comparison of the proposed protocol with other handover authentication protocols is presented in Section 5. Finally, the paper concludes in Section 6.

2. Related Work

IEEE 802.16e standard specifies three types of handover: Hard Handover (HHO), Macro Diversity Handover (MDHO) and Fast Base Station Switching (FBSS) [4]. In HHO, the MB communicates only with one BS; therefore, before connecting to a new BS, all connections with the old BS must be broken. This type of handover is simple but it is characterized by its large latency. Both MDHO and FBSS are known as soft handover where the MS could communicate with more than one BS at a time. Both the MS and BS maintain a list of BSs involved in the handover procedure. The solution of the handover problem could be classified into two categories: pre-authentication and re-authentication. In pre-authentication handover [8, 9 and 10], MSs can pre-authenticate to the most likely BSs in order to reduce the handover authentication time. Although this solution solves the latency problem, it requires a large computation and communication overhead which results from the exchange of unnecessary keys between the MS and the BSs that the MS never roams to.

In literature, many solutions have been proposed to solve the re-authentication problem. These solutions are divided into two categories [11]: public key-based schemes and symmetric key-based schemes. The public key-based schemes [12] are divided into: schemes based on ID-based cryptography [13], schemes based on blind signature schemes [14] and schemes based on capabilities [15]. These schemes suffer from a high computation cost. On the other hand, symmetric key-based schemes are divided into three categories: key hierarchy, credential tickets and Security Context Transfer (SCT) mechanisms. Key hierarchy scheme is based on using key hierarchy designed especially for handover purposes [16 and 17]. For example, the HandOver KEYing (HOKEY) [17] designed by Internet Engineer Task Force (IETF) uses EAP initiate/finish re-authentication exchange to derive a new re-authentication master key. This solution reduces the handover latency but not to the extent that could be suitable for real time applications. On the other hand, credential tickets schemes are based on incorporating a third party to securely distribute keys (such as Kerberos protocol) [18 and 19]. These schemes require a trusted third party to issue and verify the tickets. Moreover, they may require the modification of EAP protocol. In SCT mechanisms [20, 21, 22 and 23], the security context (such as negotiated key) is transferred to a target BS using Context Transfer Protocol (CXTP). Consequently, MS does not need to contact EAP server in order to perform handover. These schemes suffer from the domino effect problem in which compromising of one BS will result in compromising other BS’s that share the same security context. All the previous schemes solve the problem for intra-domain handover, where all BS’s are in the same area. For inter-domain handover, the MS must perform the full EAP authentication protocol which leads to an impractical latency for mobile WiMAX applications. In the next section, a description of the proposed protocol is detailed.

3. Inter-Domain Handover Authentication Protocol for IEEE 802.16m Networks

The proposed protocol consists of three phases: the initial phase, the intra-domain handover authentication phase and the inter-domain handover authentication phase. Fig. 1 shows an example of IEEE 802.16m network. In this network, the Access Service Network (ASN) consists of ASN Gateway (ASN-GW) and BS. An ASN-GW controls several BSs and has the role of forwarding authentication messages between the MS and the Authentication, Authorization and Accounting (AAA) server [11]. A BS provides WiMAX radio access for the MSs authenticated by the AAA server. In this paper, it is assumed that AAA server and all ASN-GW maintain trusted relations and have established secure connections. Also, each ASN-GW and BSs served by this ASN-GW maintain trusted relations and have established secure connections. In addition, an assumption is made that all entities share two constant values: a and p which will be used to calculate the Diffie-Hellman keys. Furthermore, a list of all ASN-GW identities and BS identities are previously delivered to MS. In the following subsections, description of the proposed protocol is detailed.
3.1 The Initial Phase

When an MS first accesses the IEEE802.16m network, it performs the initial authentication as shown in Fig.2. First, a full EAP authentication is executed. After a successful authentication, both AAA server and MS share a Master Session Key (MSK) which is unique for each mobile station. Then, both MS and AAA server calculates a hash value of MSK and the identity of ASN-GW1 (ASN1ID), and AAA server forwards the hashed value \( H_1 = H(\text{MSK}, \text{ASN1ID}) \) to ASN-GW1. Then, both MS and ASN-GW1 compute a Temporary CMAC Key (TCK) using \( H_1 \) and the identity of BS1 (BS1ID). Finally, ASN-GW1 forwards the temporary key TCK = \( H(H(\text{MSK},\text{ASN1ID}), \text{BS1ID}) \) to BS1. Next, BS1 sends to the MS a message (MSG-1) containing: a timestamp \( T_{BS1} \), the identity of BS1 (BS1ID) and its Diffie-Hellman component \( a^x \mod p \), where \( x \) is the Diffie-Hellman secret of BS1. Each entity stores and maintains the secrecy of its Diffie-Hellman component. To maintain message authenticity, a CMAC value of the abovementioned message is appended. After receiving MSG-1, MS verifies the freshness of the message using \( T_{BS1} \), then uses TCK to validate the CMAC value. If the CMAC value is valid, MS sends to BS1 a message containing: a timestamp \( T_{MS} \), its identity (MSID), its Diffie-Hellman component \( a^y \mod p \), where \( y \) is the Diffie-Hellman secret of MS. Again, a CMAC value of the abovementioned message is appended. Then, it calculates the Authentication Key (AK), where \( AK = a^{xy} \mod p \). After receiving MSG-2, BS1 verifies the freshness of the message using \( T_{MS} \), then uses TCK to validate the CMAC value. If the CMAC value is valid, it calculates \( AK = a^{xy} \mod p \). Then, both BS1 and MS derive the Traffic Encryption Keys (TEK) and CMAC keys as stated in [24]. The use of Diffie-Hellman method preserves the confidentiality of the derived keys which means that only MS and BS1 know the calculated key.

3.2 The Intra-Domain Handover Authentication Phase

When an MS moves from one base station to another in the same region, it sends to ASN-GW1 a request containing its identity and the identity of the target base station, for example BS2 and appends the CMAC value of the message using TCK. The CMAC value is used to avoid man in the middle and denial of service attacks. Next, both ASN-GW1 and MS compute another temporary CMAC key (TCK*) using the hash function of \( H_1 \) (hash value of MSK and ASN-GW1’s identity) and the identity of BS2 (BS2ID). Next, ASN-GW1 forwards the temporary key TCK* to BS2. Next, BS2 sends to the MS a message (MSG-3) containing: a timestamp \( T_{BS2} \), the identity of BS2 (BS2ID) and its Diffie-Hellman component \( a^z \mod p \), where \( z \) is the Diffie-Hellman secret of BS2. To maintain message authenticity, a CMAC value of the abovementioned message is appended. After receiving MSG-3, MS verifies the freshness of the message using \( T_{BS2} \), then uses TCK* to validate the CMAC value. If the CMAC value is valid, MS sends to BS2 a message
Compute $TCK = H(H(MSK, ASN1ID), BS1ID)$

MSG-1: $(T_{BS1}, BS1ID, a^x \mod p)(CMAC(TCK))$

Verify CMAC value

Calculate $AK = a^y \mod p$ and derive TEKs and CMAC keys

Fig. 2 Initial phase.

Compute $TCK' = H(H(MSK, ASN1ID), BS2ID)$

MSG-3: $(T_{BS2}, BS2ID, a' \mod p)(CMAC(TCK'))$

Verify CMAC value

Calculate $AK' = a'^z \mod p$ and derive TEKs and CMAC keys

Fig. 3 Intra-domain handover phase.
(MSG-4) containing: a timestamp $T_{MS}$, its identity (MSID), its Diffie-Hellman component $a^y \bmod p$. Again, a CMAC value of the abovementioned message is appended. Then, it calculates the Authentication Key (AK'), where $AK' = a^y \bmod p$. After receiving MSG-4, BS2 verifies the freshness of the message using $T_{MS}$, then uses TCK' to validate the CMAC value. If the CMAC value is valid, it calculates $AK' = a^y \bmod p$. Then, both BS1 and MS derive the traffic encryption keys and CMAC keys. The abovementioned steps are illustrated in Fig. 3.

3.3 The Inter-Domain Handover Authentication Phase

When an MS moves from one base station to another in a different region, it sends to AAA server a request containing its identity and the identity of the target base station, for example BS3 and appends the CMAC value of the message using MSK. Then, both MS and AAA server calculates a hash value of MSK and the identity of ASN-GW2 (ASN2ID), and AAA server forwards the hashed value $H_2 = H(MSK, ASN2ID)$ to ASN-GW2. Then, both MS and ASN-GW2 compute a Temporary CMAC Key (TCK") using $H_2$ and the identity of BS3 (BS3ID). Finally, ASN-GW2 forwards the temporary key TCK" = $H(H(MSK, ASN2ID), BS3ID)$ to BS3. Next, BS3 sends to the MS a message (MSG-5) containing: a timestamp $T_{MSG}$, its identity (MSID), its Diffie-Hellman component $a^m \bmod p$, where $m$ is the Diffie-Hellman secret of BS3. To maintain message authenticity, a CMAC value of the abovementioned message is appended. After receiving MSG-5, MS verifies the freshness of the message using $T_{BS3}$, then uses TCK" to validate the CMAC value. If the CMAC value is valid, MS sends to the BS a message (MSG-6) containing: a timestamp $T_{MS}$, its identity (MSID), its Diffie-Hellman component $a^m \bmod p$, where $y$ is the Diffie-Hellman secret of MS. Again, a CMAC value of the abovementioned message is appended. Then, it calculates the Authentication Key (AK'"), where $AK'' = a^ym \bmod p$. After receiving MSG-6, BS3 verifies the freshness of the message using $T_{MS}$, then uses TCK" to validate the CMAC value. If the CMAC value is valid, it calculates $AK'' = a^ym \bmod p$. Then, both BS3 and MS derive the Traffic Encryption Keys (TEK) and CMAC keys. The abovementioned steps are illustrated in Fig. 4. In the next section, logical analysis of the proposed protocol will be presented.

4. Logical Analysis of the Proposed Handover Protocol

BAN logic [25] is a set of rules for defining and analyzing information exchange protocols. Specifically, BAN logic helps its users to determine whether exchanged information is trustworthy, secured against eavesdropping, or both. For a successful verification of the protocol, the belief state of communicating parties should satisfy the protocol goals. The goals of the proposed protocol are: the BS and the MS believe that they share a common key and also each of them should believe that the other participant also believes in the same key.

First, the basic rules of the BAN logic are listed below:

- **The interpretation rule**
  \[
  P \models (Q \Rightarrow (X, Y)) \Rightarrow \begin{cases}
  P \models (Q \Rightarrow X), & P \models (Q \Rightarrow Y) \\
  P \models (Q \Rightarrow X), & P \models (Q \Rightarrow Y) 
  \end{cases}
  \]
  The above rule means that if P believes that Q once said a message containing both X and Y, therefore it believes that Q once said each statement separately.

- **Message Meaning Rule**
  \[
  P = P \Leftrightarrow Q, P \ll X > Y, P \neq Q
  \]
  This means that if P believes Y is a shared secret between it and Q, and P sees a message X combined with Y, this implies that P believes that Q once said X.

- **Nonce Verification Rule**
  \[
  P \models #(X), P \models Q \Rightarrow X \\
  P \models Q \Rightarrow X
  \]
  The above rule means that if P believes that X is a recent message and Q once said X, therefore it believes that Q believes in X.

- **Jurisdiction Rule**
  \[
  P \models Q \Rightarrow X, P \models Q \models X
  \]
  This rule means that if P believes that Q has jurisdiction over X, and P believes that Q believes in X, then P believes in X.

- **Freshness Rule**
  \[
  P \models #(X, Y) \\
  P \models #(X)
  \]
  The above rule means that if P believes in the freshness of X and Y, therefore it believes in the freshness of each statement separately.
Diffie-Hellman rule

\[ P \rightarrow (\alpha^x \mod p) \rightarrow (\alpha^{XY} \mod p) \]

This rule was added to the list of rules in order to complete the logical analysis of the proposed protocol. This rule means that if P believes that x is its Diffie-Hellman secret and it believes that \( \alpha^x \mod p \) is the Diffie-Hellman component of Q, therefore it believes that \( \alpha^{XY} \mod p \) is a shared symmetric key between itself and Q.

The analysis is undertaken for the initial phase only, for intra-domain and inter-domain handover authentication phases, the same analysis could be carried out. Since, the authentication between the MS and the network is performed in the EAP protocol, therefore, the authentication is considered completed between the BS1 and the MS if the following goals are achieved:

Goal 1: BS1 \( \equiv \) MS \( \leftrightarrow \) BS1

Goal 2: MS \( \equiv \) MS \( \leftrightarrow \) BS1

In order to complete the analysis, the following assumptions are made:

\[ TCK \]

\[ BS1 \equiv MS \leftrightarrow BS1 \] (1)
Eqs. (1 and 2) indicate that both MS and BS1 believe that TCK is a shared secret between them. Then, Eqs. (3 and 4) indicate that both BS1 and MS believe in the freshness of $T_{MS}$ and $T_{BS1}$ respectively. Eqs. (5 and 6) indicate that each entity believes in its Diffie-Hellman secret. After making the assumptions, the messages transferred in the initial phase are transformed into logical formulas. Finally, the basic rules of the BAN logic will be applied to the logical formulas. Following is the transformation of the proposed protocol into logical formulas:

$$a^x \text{ modp}$$

$$\text{BS1} \equiv a^x \text{ modp}$$

$$\text{MS} \equiv a^y \text{ modp}$$

$$\text{MS} \equiv (\text{BS1} \Rightarrow a^x \text{ modp})$$

$$\text{BS1} \equiv a^y \text{ modp}$$

But, MS believes in the freshness of $T_{BS1}$ (Eq.(4)). Thus, applying nonce verification rule, the following is obtained:

$$a^x \text{ modp}$$

$$\text{BS1} \equiv \text{BS1} \Rightarrow a^x \text{ modp}$$

Then, by applying jurisdiction rule using Eq. (7), the following is obtained:

$$a^x \text{ modp}$$

$$\text{MS} \equiv \text{BS1} \Rightarrow a^x \text{ modp}$$

From Equation (6) and by applying the Diffie-Hellman rule, the following is obtained:

$$a^x \text{ modp}$$

$$\text{MS} \equiv \text{BS1} \Rightarrow a^x \text{ modp}$$

Similarly, for the analysis of the second message of the protocol, by applying message meaning rule to Eq. 10 and using Eq. 1, the following can be deduced the following:

$$a^y \text{ modp}$$

$$\text{BS1} \equiv \text{BS1} \Rightarrow a^y \text{ modp}$$

But, BS1 believes in the freshness of $T_{MS}$ (Eq.(3)). Thus, applying nonce verification rule, the following is obtained:

$$a^y \text{ modp}$$

$$\text{BS1} \equiv \text{BS1} \Rightarrow a^y \text{ modp}$$

Then, by applying jurisdiction rule using Eq. (8), the following is derived:

$$a^y \text{ modp}$$

$$\text{BS1} \equiv \text{BS1} \Rightarrow a^y \text{ modp}$$

From Equation (5) and by applying the Diffie-Hellman rule, the following is obtained:

$$a^y \text{ modp}$$

$$\text{BS1} \equiv \text{BS1} \Rightarrow a^y \text{ modp}$$

From Eqs. (11 and 12), one can deduce that the proposed protocol achieves the goals of authentication and key distribution without bugs or redundancies. In the next section, comparison of the proposed protocol with other handover protocols is detailed.
5. Comparison of the Proposed Protocol with IEEE 802.16m and SCT-Based Protocols

The proposed protocol is compared with IEEE 802.16m and Fu et al. [12] protocols. The comparison will be undertaken according to: communication and computation overheads. In order to undertake the communication overhead comparison, the following parameters are defined: \( T_{EAP} \) is the delay for the full EAP authentication, \( T_w \) is the transmission latency between the MS and the BS, and \( T_a \) is the transmission latency between the BS and ASN-GW or latency between ASN-GW and AAA server. According to [4], the following values could be assumed: \( T_{EAP} = 1000 \) ms, \( T_w = 15 \) ms, and \( T_a = 20 \) ms. Table 1 shows a comparison of communication overheads concerning the proposed protocol, IEEE 802.16m and Fu et al. protocols. The table shows that the proposed protocol outperforms both the IEEE 802.16m and Fu et al. protocol for both inter-domain and intra-domain handover.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>MS</th>
<th>BS</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEEE 802.16mm</td>
<td>2( T_{DOT} ) + 3( T_{CMAC} )</td>
<td>2( T_{DOT} ) + 3( T_{CMAC} )</td>
</tr>
<tr>
<td>Fu. Et al. Protocol</td>
<td>4( T_{DOT} ) + 3( T_{CMAC} ) + 2( T_{hash} ) + 2( T_{sym} )</td>
<td>4( T_{DOT} ) + 3( T_{CMAC} ) + 2( T_{hash} ) + 2( T_{sym} )</td>
</tr>
<tr>
<td>Proposed Protocol</td>
<td>3( T_{CMAC} ) + 2( T_{hash} ) + ( T_{exp} )</td>
<td>3( T_{CMAC} ) + 2( T_{hash} ) + ( T_{exp} )</td>
</tr>
</tbody>
</table>

In order to carry out the computation overhead comparison, the following parameters are defined: \( T_{CMAC} \) is the time for CMAC operation, \( T_{hash} \) is the time to perform a hash operation, \( T_{DOT} \) is the time for a DOT operation, \( T_{sym} \) is the time to perform a symmetric operation, and \( T_{exp} \) is the time to calculate a modulo exponentiation. Table 2 shows a comparison of computation overheads concerning the proposed protocol, IEEE 802.16m and Fu et al. protocol. It has to be noted that the modulo exponentiation time is the most consuming: time among the abovementioned parameters. As stated in [12], the exponentiation time could be equal to 2 ms. Therefore, the computation overhead could be negligible compared to the communication overhead. As mentioned in the previous sections, the advantages of using Diffie-Hellman protocol are: to overcome the domino effect and also that both MS and BS shared a symmetric key which is only known to both of them which is not the case in both IEEE 802.16m and Fu et al. protocols. This leads to maintain the confidentiality of messages exchanged between both MS and BS. Although the computation overhead is higher than the other protocols, the latency performance (including communication and computation overheads) is still efficient. This is a result of the dramatically decrease in the communication overhead of the proposed protocol.

6. Conclusions

Mobile WiMAX requires the re-authentication of the mobile stations as they change from one base station to another. IEEE 802.16m uses Extensible Authentication Protocol (EAP) for authentication and key management. This requires about 1000 ms, therefore, it could not support mobile WiMAX applications such as video conference. In literature, many solutions have been proposed to solve the re-authentication problem. These solutions solve only the problem of intra-domain handover (where the target base station and the serving base station are in the same domain). For inter-domain handover (where the target base station and the serving base station are in different domains), the MS must perform the full EAP authentication protocol which leads to an impractical latency for mobile WiMAX applications. In the present paper, a protocol that aims to overcome the inter-domain handover problem is proposed. The proposed protocol uses the Extensible Authentication Protocol (EAP) for authentication and key distribution. The proposed protocol is based on the use of hash functions and the Diffie-Hellman protocol to distribute the keys between mobile stations and base stations. In order to avoid the domino effect, the Diffie-Hellman components are distributed instead of the authentication key itself. The advantages of using Diffie-Hellman protocol are: to overcome the domino effect in case of BS compromise and also that both MS and BS shared a symmetric key which is only known to both of them which is not the case in both IEEE 802.16m and Fu et al. protocols. This leads to maintain the confidentiality of messages exchanged between both MS and BS. The proposed protocol is analyzed using the BAN logic to ensure that it achieves the goals of authentication.
and key distribution. The analysis shows that the proposed protocol achieves its goals without bugs or redundancies. Furthermore, the proposed protocol is compared with other handover protocols. It has been shown that, although the use of Diffie-Hellman protocol slightly increases the computation overhead, the proposed protocol still outperforms the other protocols. This is a result of the dramatically decrease in the communication overhead of the proposed protocol.

References


Heba Kamal Aslan is an Associate Professor at Electronics Research Institute, Cairo-Egypt. She received her B. Sc. degree, M. Sc. degree and Ph. D. degree in Electronics and Communications Engineering from the Faculty of Engineering, Cairo University, Egypt in 1990, 1994 and 1998 respectively. Aslan has supervised several masters and Ph. D. students in the field of computer network security. Her research interests include: Key Distribution Protocols, Logical Analysis of Protocols and Wireless Security.