Attacks and Countermeasures on 802.16: Analysis and Assessment

Constantinos Kolias, Georgios Kambourakis, and Stefanos Gritzalis

***Abstract*—The IEEE 802.16 technology, commonly referred to as WiMAX, gains momentum as an option for broadband wireless communication access. So far, several research works focus on the security of the 802.16 family of standards. In this context, the contribution of this paper is twofold. First, it provides a comprehensive taxonomy of attacks and countermeasures on 802.16. Each attack is classified based on several factors, e.g. its type, likelihood of occurrence, impact upon the system etc. and its potential is reviewed with reference to the standard. Possible countermeasures and remedies proposed for each category of attacks are also discussed to assess their effectiveness. Second, a full-scale assessment study of indicative attacks that belong to broader attack classes is conducted in an effort to better comprehend their impact on the 802.16 realm. As far as we are aware of, this is the first time an exhaustive and detailed survey of this kind is attempted.**

***Index Terms*—WiMAX, 802.16, Security, Attacks, Survey, Simulation.**

I. Introduction

THE IEEE 802.16 standard was developed to satisfy the growing demand for Broadband Wireless Access (BWA)[1]. Commonly known as the Worldwide Interoperability for Microwave Access (WiMAX), this technology may still fall short in terms of adoption rate when compared to other popular technologies such as IEEE 802.11 [2]. Nevertheless, it is obvious that the more widespread wireless technologies (such as the aforementioned 802.11) are not appropriate for outdoor BWA applications. Thus, WiMAX is expected to be a dominant technology for the Metropolitan Area Networks (MANets) in the near future. Indeed, WiMAX attempts to overcome the last mile bottleneck issue of contemporary telecommunication networks. Among its other assets the support for all-IP core network infrastructure, low latency, advanced Quality of Service (QoS) and sophisticated security prevail.

WiMAX was partially based on the Data Over Cable Service Interface, Baseline Plus Interface (DOCSIS BPI+) protocol [3] which has been originally designed for cable modems. The first version of the standard, i.e., IEEE 802.16-2001 [4] only supported point-to-multipoint (PMP) fixed wireless access between a Base Station (BS) and several registered Subscriber Stations (SS). Since IEEE 802.16-2001 operates in the 10-66 GHz frequency range, this technology required line-of-sight (LOS) communication. The next version of the standard namely, IEEE 802.16-2004 [5] extended the frequency range into the 2-11 GHz band, thus enabling nonline-of-sight (NLOS) communication. Among other improvements in this version mesh mode was introduced. Until now, the most prominent version of the standard, namely IEEE 802.16e-2005 [6] specifies numerous major improvements including the full mobility support. Subscribers are now characterized as Mobile Stations MSs (in the following we use the terms MS and SS interchangeably). This became possible as the standard employs Scalable Orthogonal Frequency Division Multiplexing (SOFDM) in the physical layer. Additionally, it supports advanced security features such as mutual authentication for both the BS and MS. 802.16j-2009 [7] added support for multihop relays. Finally, the latest version of the standard, namely 802.16m-2011 [8] (also known as WiMAX release 2), added support for data rates as high as 100 Mbps for mobile nodes and 1 Gbps for stationary users.

As always, security is an essential prerequisite for the success of every communication technology. Wireless communications are by nature more vulnerable to a number of different attacks such as man-in-the-middle, DoS and replay. In the case of WiMAX -which was initially constructed with respect to a protocol for wired environments-, the security provision is rather inefficient as wireless and wired realms enjoy very different threat models. Moreover, with each version of the standard new improvements were added but at the same time new threats emerged. For instance, lower frequencies (and as a result NLOS communication) introduced in IEEE 802.16-2004 reduce the hardware implementation complexity and the physical placement constraints for an attacker. Similarly, mesh mode is by definition insecure as it assumes the trustworthiness of all nodes of the network. Moreover, the support for mobility, facilitates aggressors to launch attacks from virtually anywhere within the network. The relay station feature, introduced in the 802.16j-2009, announces a new element on the network, but with it, another target of possible attacks is added. The latest version of the standard added improvements in the security mechanisms that include encrypted control messages and a new version of the Privacy and Key Management protocol (PKMv3). However, since this version is very recent and there are no published works to report on the robustness of its security features so far, it will be excluded from this survey. For the same reason attacks against the 802.16j amendment are left out. Nevertheless, it is true that the rest of the amendments of the standard, despite the constant security improvements, maintain

Manuscript received 17 September 2011; revised 18 January 2012.

The authors are with the Laboratory of Information and Communication Systems Security, Department of Information and Communication Systems Engineering, University of the Aegean, Karlovassi, GR-83200 Samos, Greece (e-mail: kkolias, gkamb, sgritz@aegean.gr).

Digital Object Identifier 10.1109/SURV.2012.021312.00138

[1]A list of acronyms used in this paper are summarized in [1]

a considerable number of security inefficiencies, exposing the user and the network to a significant number of threats.

Our Contribution: This work surveys the attacks and countermeasures found in the literature against the IEEE 802.16 family of standards. The focus is on attacks, feasible against the MAC layer and up to the 802.16-2009 version of the standard. Thus, attacks against the physical layer of WiMAX, like the ones described in [9] are outside of the scope of this work. Also, this work attempts to organize and classify the possible attacks rather than simply provide an overview of the documented threats like in [10]. For each of them first the vulnerability of the protocol that makes the attack feasible is presented. Next, the attack methodology is thoroughly explained and evaluated based on a number of different criteria. It is stressed that, while a risk analysis of threats in WiMAX has been attempted in the past [11], [12], [13], our work focuses on specific attacks and not generic threat categories. This work also contributes the reproduction of characteristic attacks with special interest in simulation environment. This aims in assessing the potential threat and quantifying the possible risk in a less theoretical level. Finally, proposed remedies for numerous vulnerabilities are presented and evaluated. To the best of our knowledge it is the first time an exhaustive and detailed survey of this kind is attempted.

The remainder of this paper is organized as follows: The next section describes basic concepts of WiMAX architecture including the network structure and security mechanisms. Although this paper assumes a minimum level of familiarization with the mechanisms of 802.16 by the reader, this section is necessary for reasons of completeness. Section III gives an insight, categorizes and surveys attacks against 802.16. Section IV presents solutions found in literature for dealing with different categories of WiMAX threats. Section V contains experimental results of some attacks that pose special interest. Finally, in Section VI we conclude.

II. WiMAX Architecture

A. Protocol Stack

The IEEE 802.16 protocol is organized primarily in the Physical (PHY) and the Medium Access Control (MAC) layers. The MAC layer can be further divided into three sublayers, namely the Service Specific Convergence Sub-layer (CS), the Common Part Sub-layer (CPS) and the Security Sublayer. CS is the sub-layer that communicates with higher layers to acquire network data. In the process it transforms these data into MAC Service Data Units (SDUs). The format of the CS payload itself is CS depended.

CPS provides basically the core MAC functionality being responsible for functions such as bandwidth allocation, connection establishment, and connection maintenance.

The Security Sub-layer, addresses procedures such as authentication, authorization, key establishment, distribution and management. Also, it is responsible for encryption and decryption of traffic passing from the PHY to the MAC layer and vice versa. The security mechanisms applied here will be presented in greater detail later in this section.

The physical layer allows great flexibility to service providers in matters of cell planning, cost, radio capabilities, services, and network capacity. It supports both Time Division Multiplexing (TDD) and Frequency Division Duplexing (FDD) configurations. More specifically, the UpLink (UL) channel is based on TDMA burst transmission and is divided into a number of time slots (their number may change over time) that are assigned for specific purposes such as registration, contention, and user traffic. Each burst carries MAC PDUs of variable size. The DownLink (DL) channel makes use of Time Division Multiplexing (TDM). The multiplexed data of each MS forms a single stream that is received by all MS within the same network cell.

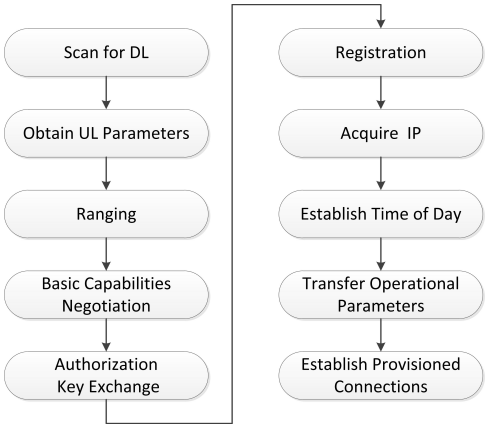


Fig. 1. Initial Network Entry

B. Network Entry

During the Initial network entry many critical parameters are negotiated between the MS and BS. From a security point of view the entire procedure is extremely receptive to violations since for most of its part the security measures contemplated by the specification have not taken place and important negotiation parameters are transmitted in cleartext. This section describes the basic steps that occur during the initial entry of an MS to the network. The overall procedure is summarized in Figure 1.

Upon initial network entry (or after loss of signal) the first action an MS does is to acquire a DL channel. The MS shall begin scanning the DL frequency band for possible channels of operation until it finds a valid DL signal. This step ends once the PHY has achieved synchronization.

To do so, at least one DL-Medium Access Protocol (DL-MAP) message must be received by the MS. The DL-MAP informs the MS about the DL-Burst Profiles. The MAC remains in synchronization as long as it keeps receiving the DL-MAP and Downlink Channel Descriptor (DCD) messages for the channel. An MS may use the information contained in the DCD to determine if the channel corresponds to its needs. Also, the MS shall search for a Uplink Channel Descriptor (UCD) message from the BS (this is transmitted periodically to all the available UL channels) for retrieving the transmission parameters of a possible UL channel.

After the MS has obtained the UL and DL parameters it will attempt to acquire the correct timing offset and make power adjustments through the process of *ranging*. The MS shall use the information contained in the UL-MAP (or UCD) message

to find an initial ranging interval. Usually, the BS allocates an initial ranging interval consisting of many Transmission Opportunities (TO). For Single Carrier (SC) and OFDM PHY, the MS shall construct an RNG-REQ message. Then the MS can transmit the RNG-REQ message in one of the known TO. Typically, there are only 3 TO in a 5 ms frame thus there is high probability of a collision to occur. To reduce collisions the 802.16 specification dictates that the nodes should pass a period of inactivity of random duration known as Backoff (BO). If a collision occurs, the MS will eventually become aware of it since the corresponding RNG-RSP message will not arrive to the device within the expiration of the T3 timer (set to 200 ms by default). The collided nodes will attempt to resend the RNG-REQ message after a random waiting time but the waiting time interval will be doubled (until a maximum value is reached). This process will repeat (as long as MSs collide) up to a defined maximum number of retries. The aforementioned process is known as Truncated Binary Exponent Backoff (TBEB).

Once the RNG-REQ message has been received by the BS, the latter will construct an RNG-RSP message and send it using the Initial Ranging Connection Identifier (CID). This message exchange shall result in the MS acquiring Basic and Primary Management CIDs as well as information about RF power level adjustment, offset frequency adjustment and timing offset corrections. Figure 2 presents the entire initial ranging procedure.

After ranging has successfully taken place, the MS will send to the BS an SS Basic Capability (SBC)-REQ message to inform it of its basic capabilities. The BS responds with an SBC-RSP message containing only the capabilities both the MS and BS can support.

Upon successful capabilities negotiation, MS authorization and key exchange follows. The details of this procedure are analysed in greater depth in the next section.

Registration is the process that takes place after successful authorization. During this step the MS gets the Secondary Management CID. This means that the MS is actually granted entry into the network. This process involves the exchange of a pair of REG-REQ and REG-RSP messages. When both messages are successfully received, the BS will authorize the MS to forward traffic to the network.

Typically, the MS shall invoke Dynamic Host Configuration Protocol (DHCP) mechanisms for receiving all relative parameters, establishing IP connectivity and obtaining an IP address. The versions of the IP that are supported by the MS are indicated in the REG-REQ message with the default value being IPv4.

Both the MS and BS need to be synchronized as the management system requires the current date and time for time-stamping logged events. For this reason, request and response messages are exchanged with a time server. The MS's secondary management connection is utilized for this process. This step is crucial for ongoing operation although not obligatory for a successful registration.

Next, the MS shall receive the MS Configuration File using Trivial File Transfer Protocol (TFTP) on the secondary management connection. This file consists of a number of configuration settings that are encoded in Type-Length-Value (TLV) format. The MS must notify the BS of successful receipt of this message by transmitting a Configuration File TFTP Complete (TFTP-CPLT) message on the primary management connection.

As a final step, the BS shall send several Dynamic Service Addition (DSA)-REQ messages to the MS for creating new service flows. The MS responds with DSA-RSP messages.

C. Security Architecture

IEEE 802.16 relies on the Security Sub-layer to provide security to the end-subscriber and the network. This part of the protocol is where all the necessary cryptographic transformations are applied to the MAC PDUs. This is necessary to provide: (a) privacy, confidentiality and authentication to the subscribers, and (b) protection from theft of service to the service providers.

The basic mechanism of security enforcement in 802.16 is the Privacy Key Management (PKM) protocol. Both 802.16-2009 and 802.16e-2005 support two versions of the PKM protocol. Mainly, PKM is responsible for authorization of subscribers and distribution of the keying material to the MS. Secondarily, it controls the application of the negotiated encryption algorithms to the data traffic. Actually, the PKM tasks can be divided into three distinct undertakings namely, authentication, key exchange, and encryption with a brief step for the key derivation that takes place in between the authorization and key exchange phases.

1) Authorization: The step of authorization happens first in the PKM protocol. The messages exchanged in this step differ for the two versions of the protocol. In PKMv1 the authorization process is initiated by the Authentication Information message which is sent by the MS to the BS. This message contains the MS manufacturer's X.509 certificate, (the manufacturer may have issued itself this certificate) and it is strictly informative.

This message is then followed by an Authorization Request message sent again by the MS to the BS. This is comprised of the following information: (a) The manufacturer-issued X.509 certificate of the MS, (b) a description of the cryptographic capabilities the MS supports, and (c) the MSs Basic CID. The purpose of this message is to request an Authorization Key (AK), and at the same time be assigned with the Security Association Identifiers (SAID) (matching the corresponding Static SAs) that the client has the right to participate in.

In response, the BS validates the MS's identity and determines the encryption algorithm (among the commonly supported ones), activates an AK for the MS, and constructs an Authorization Reply message. The latter is sent to the MS and it consists of the following fields: (a) the active (for this particular MS) AK which is encrypted with that MSs public key, (b) a 4-bit key sequence number used to distinguish between successive generations of AKs, (c) the lifetime of this key, and (d) the identities (i.e., the SAIDs) and properties of the Primary and Static SAs for which the MS is authorized to obtain keying information. Protocol 1 illustrates the flow of messages and their contents during this step.

It is obvious that authentication in PKMv1 is one way, meaning that the BS can authenticate the MS but not vice-versa. This introduces a vulnerability that soon became the

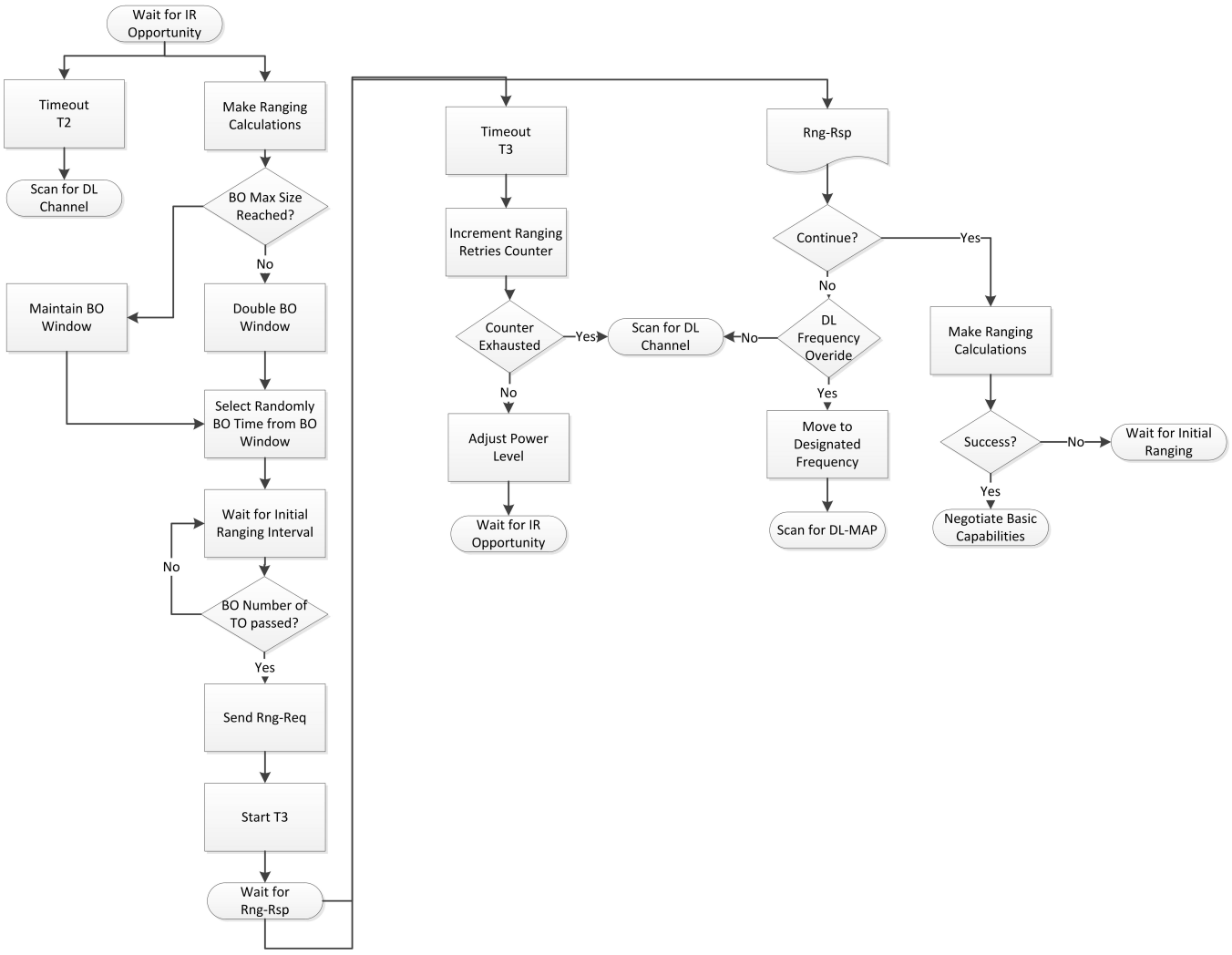


Fig. 2. Flowchart of the Initial Ranging process

Protocol 1 Authorization Step of PKMv1 Protocol

MS $\rightarrow$ BS: Authentication Information($Cert_{Manufacturer}$)
MS $\rightarrow$ BS: Authorization Request($Cert_{MS}$, [Cryptographic Capabilities], Basic CID)
BS $\rightarrow$ MS: Authorization Reply($Enc(AK)_{PK}$, Sequence Number, Key Lifetime, [SAIDs])

cause of many attacks. As a result, PKMv2 was introduced and the authorization part is slightly modified to support mutual authentication. As in PKMv1, the second version of the PKM protocol dictates that the authorization process must start with the transmission of the informative message Authentication Information. This message is the same as in PKMv1.

Following the Authentication Information the MS must issue an Authorization Request. The format of this message is modified in PKMv2. The Authorization Request includes: (a) The manufacturer-issued X.509 certificate of the MS, (b) a description of the cryptographic algorithms supported by the MS, (c) the SSs Basic CID, and (d) 64-bit random number generated by the MS. The last field is the only new one added since the original PKM protocol to the Authorization Request message.

In response, a BS sends back an Authorization Reply message. This message has been rectified more extensively and its fields now include: (a) The BSs X.509 certificate; It is used to verify the BSs identity and to guarantee the authenticity of this message, (b) the pre-PAK key which is encrypted with the MSs public key; Only the owner of the corresponding private key will be able to decrypt it, (c) a 4-bit PAK sequence number, used to distinguish between successive generations of AKs, (d) the PAK lifetime, (e) the identities (i.e., the SAIDs) and properties of the SAs for which the MS is authorized to have keying material, (f) the 64-bit random number generated by the MS, originally contained in the Authorization Request message; This field is included to ensure that the Authorization Reply corresponds to the correct Authorization Request message. Additional fields are: (g) a new 64-bit random number generated by the BS and (h) the RSA signature over the entire message. This allows the MS to verify that the BS is indeed the author of the Authorization Reply message. This process is performed using with the public key of the BS which is acquired by the certificate contained in the message. In other words, this extra field allows for mutual authentication.

Following the message above the MS replies with an Authorization Acknowledgement message (or an Authentication Reject message in the case where the BS will reject the MS). The Authorization Acknowledgement includes: (a) the 64-bit random number originally contained in the PKMv2 RSA-Reply message, (b) an Authentication result code which can be "success" or "failure", (c) an Error code which indicates the reason for the reject, (d) an optional Display string which includes a phrase for the reason for the reject rather than just a code number, (e) an RSA signature over the entire message.

Protocol 2 Authorization Step of PKMv2 Protocol

MS $\rightarrow$ BS: Authentication Information($Cert_{Manufacturer}$)
MS $\rightarrow$ BS: Authorization Request(Rad_{MS}, $Cert_{MS}$, [Cryptographic Capabilities], Basic CID, $Signature_{MS}$)
BS $\rightarrow$ MS: Authorization Reply($Cert_{BS}$, $Enc(pre-PAK)_{PK_{MS}}$, Sequence Number, PAK Lifetime, [SAIDs],Rnd_{MS}, Rnd_{BS}, $Signature_{BS}$)
MS $\rightarrow$ BS: Authorization Acknowledgement (Rnd_{BS}, Result Code, Error Code, Display String, $Signature_{MS}$)

2) Key Derivation: After the PKM authentication phase, normally the MS is in possession of some keying material. A small step where the MS derives the appropriate keying material takes place before the PKM can proceed to the next phase. The key derivation process is different between the two versions of the PKM protocol. Actually, the keys in 802.16 form a hierarchy. A key of a higher level is used to produce the key of the immediately lower level. All key generations in PKMv2 are produced using the Dot16KDF function. This function takes 3 arguments: (a) A keying material of a higher level, (b) a string used to alter the output of the algorithm, and (c) a number used to indicate the length of the generated key.

More specifically, the RSA-based authorization process results in the creation of the pre-Primary Authentication Key (pre-PAK) while the Extensible Authentication Protocol (EAP) based authentication process produces the Master Session Key (MSK). These two keys constitute the basis of all other keying material and they are placed in the top of the key hierarchy. In RSA-based authorization a pre-PAK is used to generate the Primary Authentication Key (PAK). Optionally, the EAP Integrity Key (EIK) can also be generated from the pre-PAK. EIK is used for transmitting authenticated EAP payload. In EAP-based authorization the 512 bits MSK, is simply truncated to 160 bits to derive the Pairwise Master Key (PMK). One of the PAK, PMK or both (according to the authentication method that was used) will be provided as input to the Dot16KDF function to produce the AK. The Key Encryption Key (KEK) is derived directly from the AK.

Message Authentication Code (MAC) keys are used to sign management messages. This procedure is performed to guarantee the authenticity of these messages. The IEEE 802.16 supports two MAC modes namely Cipher-based MAC (CMAC) and Hashed MAC (HMAC). The one to be used is negotiated during the MS Basic Capabilities negotiation phase. Different MAC keys exist for UL and DL messages. The Cipher-based MAC Key for Uplink (CMAC_KEY_U) is used for signing messages in the uplink while the Cipher-based MAC Key for Downlink (CMAC_KEY_D) is used for the same purpose in the downlink. This only applies for the cipher-based MAC mode. For the hash-based MAC mode corresponding keys exist, i.e. HMAC_KEY_U, HMAC_KEY_D). In the case of HMAC these keys are derived directly from Dot16KDF function while in the case of CMAC and for versions later than 802.16e a corresponding prekey is generated first (look figure 3 for more details). Also there are different keys for broadcast and unicast messages. In any case, MAC keys are derived directly from AK by simply using different string and key size arguments in the Dot16KDF function for each mode. Figure 3 depicts a diagram of the complete key derivation flow.

3) Handshake: The next phase is a three way handshake. It's main role is to confirm that both the MS and BS have indeed the correct AK from the previous procedure. Additionally, the handshake protocol takes care of secondary procedures such as key activation and SA parameters negotiation, security parameters confirmation etc. For this purpose the BS shall send an PKMv2 SA-TEK-Challenge which simply includes: (a) a random number, (b) sequence number for the new AK, (c) the ID of the new AK, (d) Key Lifetime all protected by (e) the HMAC/CMAC.

The SS shall respond with a PKMv2 SA-TEK-Request to the BS. This message includes: (a) the random number the MS received from the PKMv2 SA-TEK-Challenge, (b) a random number the MS produces, (c) the sequence number of the new AK, (d) the ID of the AK, (e) the cryptographic suites supported by the MS, (f) the security capabilities of the MS, and (g) the HMAC/CMAC of the entire message.

Upon reception of PKMv2 SA-TEK-Request, a BS first confirms that the AKID contained in the message refers to a valid AK and then it verifies the HMAC/CMAC. After that the BS will check if the random value sent matches the one contained in the PKMv2 SA-TEK-Request. If any of the three aforementioned tests fails the BS will simply ignore the message. Finally, the BS will make sure that the security capabilities encoded in the Security Negotiation Parameters attribute are the same with the security capabilities provided by the MS through the SBC-REQ message. If not, the BS should report the inconsistency to higher layers but might as well accept the message. If the validation of the PKMv2 SA-TEK-Request is successful, the BS shall send a PKMv2 SA-TEK-Response back to the SS. This message includes all the fields of the PKMv2 SA-TEK-Request message plus a TLV list of the SAs, their identifiers (SAID) any additional properties of the SA (e.g., type, cryptographic suite) that the SS is granted access to. The TEK-Parameters attribute in that list contains keying material such as the TEKs remaining key lifetime, its key sequence number and the Cipher Block Channing Initialization Vector (CBC IV). The HMAC/CMAC is the last field of this message.

Protocol 3 Three Way Handshake of PKMv2 Protocol

BS $\rightarrow$ MS PKMv2 SA-TEK-Challenge(Rnd_{BS}, AK Sequence Number, AKID, Key Lifetime, HMAC/CMAC)
MS $\rightarrow$ BS PKMv2 SA-TEK-Request(Rnd_{BS}, Rnd_{MS}, AK Sequence Number, AKID, Security Capabilities, Security Negotiation Parameters, HMAC/CMAC)
BS $\rightarrow$ MS PKMv2 SA-TEK-Response(Rnd_{BS}, Rnd_{MS}, AK Sequence Number, AKID, SA TEK Update, Frame Number, SA-Descriptor, Security Negotiation Parameters, HMAC/CMAC)

4) TEK Transportation: As already mentioned, TEK is responsible for the encryption of traffic. The BS alone is

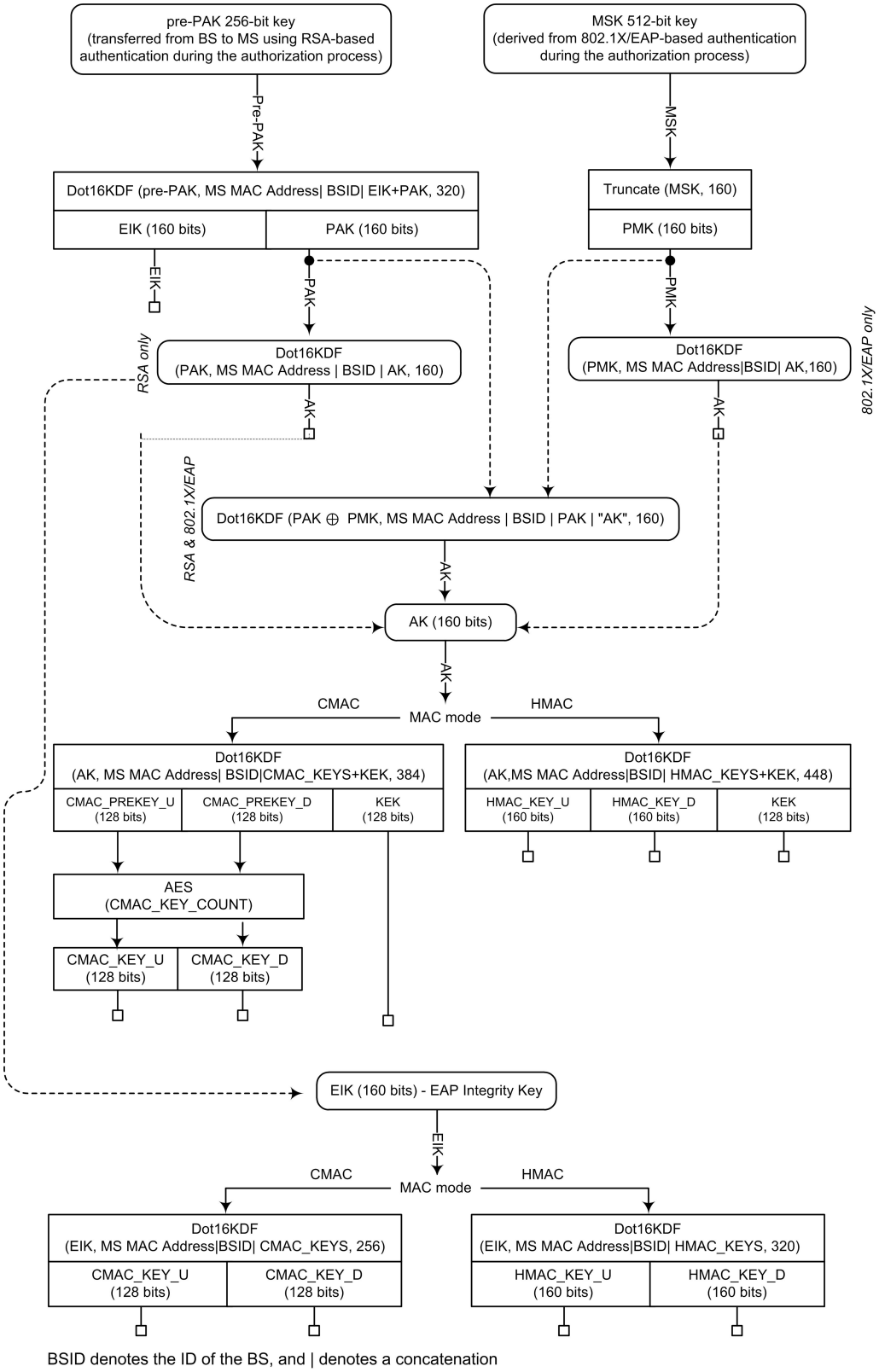


Fig. 3. Complete Message Exchange and Key Derivation

in charge for the creation of this key, thus it must securely transmit it to the MS. The pair of *PKM-REQ: Key Request* and *PKM-REP: Key Reply* messages exist for this purpose. PKM-REQ is comprised of the following fields: (a) The Key Sequence Number, which allows the BS to determine the AK used for the production of the corresponding UL HMAC/CMAC Key, (b) the ID of the SA whose TEK is requested, and (c) the HMAC/CMAC digest over the entire PKM-REQ message payload.

After verifying the authenticity of the message the BS responds with a PKM-REP message. The fields of this message are the following: (a) Key Sequence Number, (b) SAID, (c)

TEK-Parameters (Older), (d) TEK, Key Lifetime, (e) Key Sequence Number, (f) CBC-IV, (g) TEK-Parameters (Newer), (h) TEK, Key Lifetime, (i) Key Sequence Number, (j) CBC-IV, and (k) HMAC/CMAC digest over the entire message payload.

It is to be noted that a unique state machine is maintained by the MS for each SAID contained in the PKM-RSP message. Each state machine is responsible for the initial establishment of TEK as well the periodic refreshing of those keys.

5) Traffic Encryption: After successful TEK exchange, both the MS and BS are able to encrypt/decrypt traffic, using this key. Note that the generic MAC header is not included in the encryption. Multiple encryption algorithms are supported. When DES algorithm in CBC mode is used, the CBC IV for the DL, shall be calculated by performing the XOR function to IV parameter included in the PKM-REP message and the current frame number. For the the UL, the CBC IV shall be calculated by performing the XOR function to IV parameter included in the PKM-REP message and the number of the frame where the relevant UL-MAP has been transmitted. If the AES algorithm in CCM mode is used then the MAC PDU payload shall always be prepended with a 4-byte packet number which will never be encrypted. Also, the MAC PDU shall be appended an 8-byte integrity check value which will be included in the encryption. Last, if AES in CBC mode is used then the CBC IV is calculated as the result of the the IV parameter included in the PKM-RSP message XORed with the concatenation of: (a) the 48-bit MAC PDU header, (b) a 32-bit PHY Synchronization value of the MAP that a data transmission occurs, and (c) the XOR value of the 48-bit MS's MAC address and the Zero Hit Counter. The complete sequence of messages exchanged during the PKMv2 protocol is illustrated in figure 4.

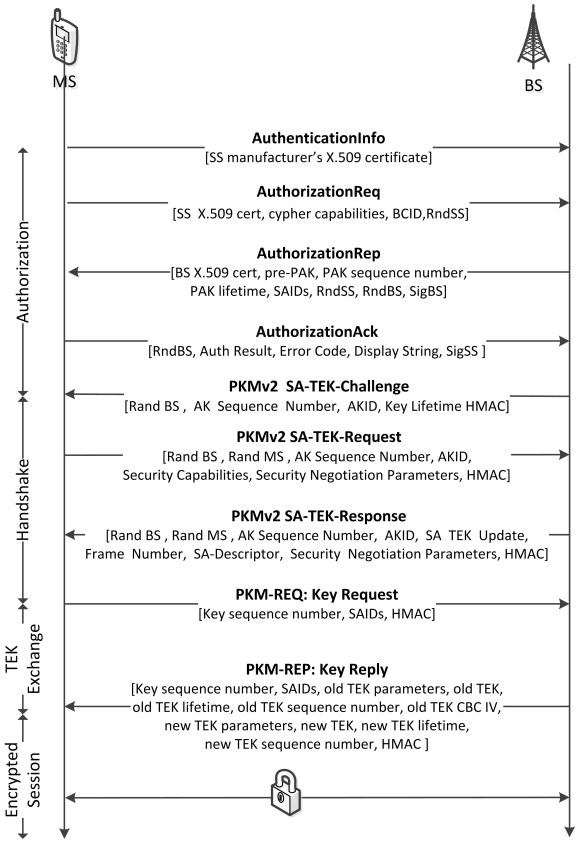


Fig. 4. PKMv2 Phases and Messages

III. Survey of Attacks

In this section we describe existing attacks against the IEEE 802.16 found in literature. The attacks are organized based on their type into several classes. Note that many of the attacks described in the literature are theoretic and we argue that some of them are not feasible under realistic conditions, although the specification does leave a margin for a security breach. Due to the lack of attack naming, we have adopted a custom naming convention to assist the readability of the paper. This naming convention follows the pattern "Message Effect" whenever this is possible. There are 2 cases where the attacks are already named in the original papers so that name was favoured and maintained in this work. A risk analysis for each one of the described attacks is also appended. Table I gathers and presents the attacks that will be described hereunder. We analyse and evaluate the severity of each attack according to a modified version of the methodology presented in [11], [13] (which in turn is a stricter version of a methodology developed by ETSI [14]). Specifically, we classify attacks according to the risk they impose to the studied system as: *Major*, *Moderate*, *Minor*. This classification is done by taking into account two factors:

1) *Likehood of occurrence* - This criterion indicates the possibility of an attack to be implemented by exploiting vulnerabilities of the system. The attack is considered unlikely if its implementation cost is high, major technical obstacles exist, or the risk of the attacker to be exposed is high. An attack is possible if the cost of the attack as well as the risk of exposing herself are moderate and the technical difficulties are solvable. An attack is likely if the associated costs and risks for the attacker are low and there are no technical difficulties associated with the attack.
2) *Impact upon the system* - This criterion is an indicator for the possible consequences to the system, provided that the attack succeeds. The attack is considered to have low impact if it affects small number of users, for a short amount of time and simply generates commotion to the system. An attack is considered of medium impact if it succeeds to afflict loss of service, affects a larger number of users but still its consequences are reversible. An attack is considered of high impact if it affects a large number of users for a significant amount of time and causes financial losses for the provider or loss of privacy/confidentiality for a user.

To formalize the aforementioned model we assume the following sets:

$C = \{Ex, Ma, In\}$ represents the Cost of the attacker with Ex being Expensive, Ma being Manageable, and In being Inexpensive.

$D = \{Ha, So, Ea\}$ represents the Difficulty to implement the attack with Ha being Hard, So being Solvable, and Ea being Easy.

$R = \{Hi, Mo, Lo\}$ represents the Risk for the attacker associated with this attack Hi being High, Mo being Moderate, and Lo being Low.

$T = \{Sh, Ln\}$ represents the Time span of the attack with Sh being Short, and Ln being Long.

$S = \{Sm, Me, La\}$ represents the Population of users affected by the attack with Sm being Small, Me being Medium, and La being Large.

$O = \{A, DoS, LoP, ToS\}$ represents the Outcome of the attack with An being Annoyance, DoS being Denial of Service, LoP Loss of Privacy and ToS being Theft of Service. Table I contains an evaluation of all the attacks that are going to be discussed in the process of this work according to this model. For all possible Threat, Likehood and Impact values with respect to the aforementioned characteristics the reader should consult the appendix.

A. Ranging Attacks

As already mentioned, one of the basic steps of initial network entry is ranging. This procedure aims in both BS and MS acquiring the correct timing offset and making the correct power adjustments so that their transmissions are aligned for the chosen physical method. When a ranging transmission opportunity occurs (this info is contained in the UL-MAP message) for the first time, the MS shall send an RNG-REQ message. Once the BS receives a decodable RNG-REQ message, it shall assign Basic and Primary Management CIDs for the MS and commit bandwidth. At the same time, the BS shall calculate any Radio-Frequency (RF) power level, frequency offset as well as any timing offset adjustments necessary for optimal communication. Finally, it will construct an RNG-RSP message with all this information and transmit it using the Initial Ranging CID. These two messages are responsible for the ranging procedure and have similar format. In more detail the RNG-RSP message defines: (a) The Basic and Primary Management CIDs for this SS, (b) Information about RF power level adjustment, (c) Information about offset frequency adjustment, and (d) Information about timing offset corrections.

This information is encoded as TLV fields and since not all fields are mandatory this message has a variable length. If the status of the RNG-RSP message is "success", the initial ranging procedure shall terminate. In case where the Ranging Status field is "continue" the Basic CID shall be used and MS and BS shall continue exchanging RNG-REQ and RNG-RSP messages for fine-tuning the parameters mentioned above. Once the RNG-REQ is within the tolerance threshold of the BS, the MS shall join data traffic in the UL. If the Ranging Status is "abort" then the MS repeats the cycle of initial network entry by scanning for DL frequency. Besides initial network entry, ranging also occurs at predefined time intervals. Periodic ranging allows the MS to adjust its transmission parameters so that it can maintain optimal UL communication with the BS. Periodic ranging may be also initiated by the BS. For this reason the MS should always be able to accept RNG-RSP messages in an unsolicited manner. These messages are not encrypted or integrity protected and they are stateless, i.e. an MS will proceed to actions dictated in an RNG-RSP message if that message is addressed to it and appears to be well-formed. Whatever the case, an attacker may manipulate the ranging messages in many ways to affect single users or the entire network. The relevant attacks found in literature are analyzed below.

1) RNG-RSP DoS Attack: This attack can be addressed to a single target MS or multiple ones. Its methodology is simple enough but the challenge in implementation lies in its high cost. In both cases the aggressor must know the radio channel of the network to be attacked and possess a BS like equipment that will allow him to transmit RNG-RSP messages. Recall that RNG-RSP messages can be transmitted in an unsolicited manner which makes the attack possible. In the first case, the attacker must also know the CID used by the victim MS. This information can easily be sniffed from any (unencrypted) management message exchanged between that specific MS and the BS. After that, the attacker forges an RNG-RSP message with the "Ranging Status" field set to "abort" and sends it to the victim [15], [16]. This will force the user/victim to disconnect from the network immediately. In the second case, the attacker simply needs to cycle through all 65,536 possible CIDs in a brute force manner, and send one forged RNG-RSP message for each CID. After that the victim device will attempt to reconnect to the network by executing Initial Network entry. Up to this point the effect of this attack is simply annoyance. By repeating this procedure the aggressor can achieve DoS to an even larger number of users, since each attack round forces the MS to a heavy signalling procedures (i.e. Initial Network entry). The DoS effect will continue only as long as the target MS(s) remain inside the influence field of the attacker and she actively and continuously transmits bogus RNG-RSP messages. Take into account that this also increases the risk for the attacker of being detected. A serious attacker might choose to collaborate with others or be on the move to avoid detection. Therefore, this attack is classified as major.

2) RNG-RSP Downgrading Attack: An RNG-RSP message can be manipulated in a number of different ways by an attacker aiming to disrupt the normal MS communication. For example, the attacker may alter the RNG-RSP frequency field in order to force the victim MS to shift to another channel. In a typical scenario, the MS will have to rescan many frequencies (wasting 5 ms in each one) until it finds the proper channel. As a consequence, this would disrupt the proper operation of the MS. Similar results can be achieved by shifting only the UL or the DL channel, or by altering other fields of the same message such as the Timing Adjust and Power Level Adjust [17].

This attack is quite similar to the RNG-RSP DoS Attack in matters of implementation methodology, implementation difficulty, realization cost, risk and possible effects. Again the attacker can hope for annoyance or even DoS if she is willing to take the analogous risks. Actually, the authors [18] implemented this attack using experimental equipment to demonstrate that when the power level field of the RNG-RSP message is set to minimum, then loss of service (annoyance) occurred for approximately 10 sec. Because the attack methodology is not trivial while there are high costs and risks associated, this attack should be considered a minor one. One interesting modification transforms it to a stepping stone for

unleashing more dangerous attacks. For example, a potentially more dangerous scenario would be to shift the victim MS to a frequency where a rogue BS set by the attacker exists.

3) RNG-RSP Water Torture Attack: This is a modification to RNG-RSP Downgrading Attack but with totally different focus on the desired effects. An attacker might forge and send an RNG-RSP message with the Power Level Adjust field set to the maximum value. This can force the MS to constantly operate in higher energy requirement state, thus causing a quicker drain of its battery resources [17].

Similarly to the already described RNG-RSP Downgrading Attack, this one is easy to implement but has high implementation costs due to its requirements for a special BS like equipment. The two attacks are also similar in methodology. The effects, which are higher battery depletion rate persist for a considerable time. That is until next ranging/periodic ranging. Thus, the risk for the attacker is smaller since she can simply transmit the message and move. On the other hand, the effects of the attack, namely the drain of energy resources are not expected to be dramatic so the attacker can hope for simple annoyance and nothing more. Therefore, an RNG-RSP Water Torture Attack is unlikely to lure serious attackers. Considering all the above, this attack is classified as minor.

It is to be noted that watertorture attacks are generally considered as physical level attacks. However, despite the fact that this survey focuses on MAC layer attacks, we have chosen to include watertorture attacks in the analysis because the attacker has to exploit vulnerabilities on the MAC plain.

4) RNG-REQ Downgrading Attack: The RNG-REQ message can also be used to inform the BS about the preferred DL burst profile. However, an attacker might take advantage of this. Specifically, by replacing the optimal burst profile with a least effective one the attacker may achieve to downgrade the service [17], [16].

The effectiveness of the attack depends on how well-manipulated the selected burst profile is. This kind of information cannot be easily deduced for any given MS. This limits the attacker to thoroughly work against a specific victim or a small number of target MSs. Also, as already mentioned, the possible outcome of a successful RNG-REQ Downgrading Attack is simply annoyance. Therefore, this attack is considered as minor.

5) RNG-REQ DDoS Attack: In this case a set of collaborating attackers may produce a large number of fake RNG-REQ messages (with different values each time) and simultaneously transmit them to the target-BS in order to waste its resources [19]. Constructing and sending multiple RNG-REQ messages with random fields and fake MAC ID, in contention mode, is not a resource intensive process for an attacker. On the other hand, the response part in the BS side is a multi-step process which consists of allocation of Basic and Primary management CIDs, deciding weather the signal is good enough or any adjustments are necessary, constructing an RNG-RSP etc.

A collaborative attack of this kind is expected to cause considerable burden in the BS which will possibly result in lower quality of service or even Distributed DoS (DDoS) for all legitimate MSs connected to the specific target-BS. Actually, when an attacker attempts such an attack she affects the system in many different ways. For example she a) artificially increases the number of collisions in the network, b) imposes burden on the BS by forcing it to conduct the ranging process for a large number of virtually non-existing MSs, and c) tricks the BS into ranging and then committing bandwidth and CIDs to fake MSs. The authors in [19] seem to have focused on the impact of this attack, to the second and third aspect mentioned previously and through simulation experiments they prove that this is a hazardous one. We argue that the most harmful aspect of this attack, is the first one but is surprisingly, neglected in literature. Due to the inefficiencies and bad mechanics of the TBEB algorithm which takes place during initial ranging, the first aspect of the attack can cause havoc to the signaling plane by transmitting extremely low volume of attack traffic. If the aggressor manages to transmit, say an RNG-REQ on every TO of each frame then she will dramatically augment the number of collisions which in turn will lead to a dramatic increase in the access delay. The same mechanism for collision avoidance is applied in the 802.11 and has been target of criticism [20] there too. The work in [21] proposes a mathematical model for the WiMAX which unlike most existing models [22], [23] for initial ranging, is not based in a stochastic Markov chain.

The only real requirement in matters of implementation methodology is the attacker to have control over a small number of programmable MSs and synchronize their actions. The cost of such devices is low. Since at this phase addressing information has not been assigned to the MS (and MAC address field contained in the RNG-REQ can easily be spoofed) the BS has not any means of recognizing the attackers. Additionally, the small attack traffic makes it extremely difficult for external tools such as Intrusion Detection Tools (IDS) to recognize such behaviour as intrusive. Therefore, this attack is classified as major.

6) MOB_ASC-REP DoS Attack: It is possible for an MS during its ranging procedure to receive an Association Result Report (MOB_ASC-REP) message instead of several RNG-RSP ones. This happens when association level 2 is used. In this case, the RNG-RSP information that is sent by each target BS is gathered to the serving BS over the backbone network. The BS then aggregates all the data from the RNG-RSP messages to a single MOB_ASC-REP message which will be transmitted over the Primary Management CID. Since the MOB_ASC-REP report messages are unprotected it is possible for an adversary to forge them stating that no services are available from all the target BSs [17], [24], [25].

A MOB_ASC-REP DoS Attack will prevent the victim MS from associating with the optimal BS which will be translated as lower QoS for the target MS. This attack assumes that the attacker transmits this kind of messages over a BS like equipment. This automatically increases the implementation costs. The attacker must also know if and when the victim MS maintains level 2 association which increases the complexity of the attack methodology. For this reason also this type of attacks are typically unleashed against a single MS or a small number of MSs. For these reasons it is less likely to be favoured as the attack strategy by a serious aggressor, therefore, it is considered minor.

B. Power Saving Attacks

The IEEE 802.16e specification introduced support for mobile devices. Since most devices of this type are battery supported, the specification included power-saving features in order to prolong the battery life of MSs. Power Saving Class (PSC) is a concept that defines a group of connections that have similar demand properties.

Three types of PSC exist: (a) Power saving class of type I which is preferred for connections of Best-Effort (BE), Non-Real Time Variable Rate (NRT-VR) type, (b) Power saving class of type II, which is recommended for Unsolicited Grant Service (UGS), Real-Time Variable Rate Service (RT-VR) connections, and (c) Power saving class of type III, which is recommended for multicast connections and management operations. In 802.16 power saving is achieved with the MS functioning in Sleep or Idle mode. Sleep mode is a state during which the MS turns off various functions thus becoming unavailable for pre-negotiated (with the serving BS) periods of time. Sleep mode is characterised by intervals of unavailability and availability in DL or UP. During an unavailability interval, the BS shall not be able to transmit to the MS (or receive from it); therefore, the MS can shutdown several physical operation components. During an availability interval, the MS is expected to receive all transmissions in the DL (or transmit in the UL) normally as in no-sleep state. During this time-frame the MS may verify synchronization with the BS and may participate in periodic ranging. Sleep mode may be initiated by the MS or the BS in a number of different ways depending on the PSC. The MS initiates Sleep mode after an inactivity timeout. To do this it transmits a MOB_SLP-REQ message or a BR and UL sleep control header. The BS shall respond to these messages with a MOB_SLP-RSP message or a DL sleep control extended sub-header, respectively. A PSC may also be activated in RNG-REQ and RNG-RSP message by the appropriate TLVs. The BS, on the other hand, may force an MS to re-activate a previously defined PSC by transmitting a MOB_SLP-RSP or DL Sleep control extended subheader in unsolicited manner. Power saving class becomes active at the frame specified as start frame number for the first sleep window. Thereafter, each sleep window is double the size of the previous one, but not greater than a maximum value. The 802.16 sleep mode process is based on a Binary Truncated Exponent algorithm (BTE) for deciding sleep intervals. This algorithm does not consider the delay of the packets and it has been the center of criticism [26]. The algorithm can be described as:

$$\begin{cases} T_{min}, \text{if } n = 1 \\ min(2^{n-1}T_{min}, T_{max}), \text{if } n > 1 \end{cases} \quad (1)$$

The BS transmits a MOB_TRF-IND with negative value on each availability interval, thus the MS repeats the sleeping cycle. During unavailability intervals, if the BS receives traffic addressed to a sleeping MS it may buffer the corresponding MAC SDUs and transmit them upon the wake. Alternatively, the BS may choose to discard them. The active state of a power saving class is terminated by the BS by sending a MOB_TRF-IND message with positive indication. This decision is made by the BS if there is traffic pending for

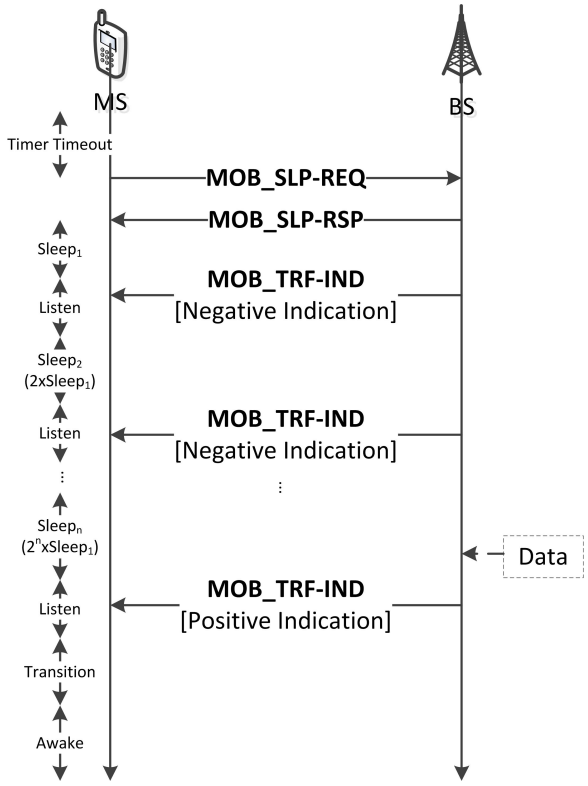


Fig. 5. Sleep Mode of Class 1 operation

that MS. In more rare situations a BS may include a positive indication even if there is no DL Traffic to be sent to the MS, given that the MSs periodic ranging is scheduled to start within the next sleep window. A power saving class can also be deactivated by MOB_SLP-REQ/BR and UL sleep control header or MOB_SLP-RSP/DL sleep control extended subheader messages. The Idle mode is a similar power-saving mechanism. It is intended mostly for situations where an inactive MS traverses a large geographic area populated by multiple BSs. Idle mode is built around the concept of paging group. A paging group is a set of BSs that maintain the same list of MSs in Idle mode and have relative configuration. This removes the need for registration at a specific BS which in turn takes away the need for handover. Upon traffic indication the MS will exit the Idle state but unlike Sleep mode it will then have to perform the network re-entry process. Both these modes aim to minimize power consumption and relieve the network from the redundant traffic of an inactive MS. Figure 5 illustrate the sleep mode procedure. Note that the implementation of both Sleep mode and Idle mode is optional for the MS but mandatory for the BS.

1) Signaling DoS Attack: Taking advantage of the power saving features the authors in [27] described an attack for the UMTS networks but indicated that it should be possible to be applied in the WiMAX realm too. According to their study in this attack an attacker could easily cause problems in a network by introducing minimal traffic to it. More specifically, the attacker would fabricate simple TCP/IP packets, for example with empty payload (40 bytes in size), and send them to several sleeping (or idle) MSs at once. In this way the attack traffic generated is as small as 64 bps and by using a

cable modem with 1.5 Mbps uplink bandwidth she can affect simultaneously 24 K MSs approximately. As data becomes available for a specific MS the BS shall have to wake it up which in turn will fall back into sleep state right after the inactivity timeout (the authors assume that this value is 5 sec for many manufacturers). By retransmitting this packet in intervals slightly larger than the inactivity timeout of the MSs the attacker will trap the network into a repeating process of "waking up-putting to sleep" many devices which can prove a great signalling burden.

Indeed the described methodology is simple enough and the cost for unleashing such attack is extremely low. However, there are several issues associated with the described methodology. First of all, the Sleep mode as well as Idle mode is optional for an MS. Moreover, it is expected that the inactivity timeout counter varies significantly among device manufactures as the default value for this parameter is not defined in the specification. The attacker will be restricted in transmitting intervals equal to the maximum of inactivity timeout value. Also, the attacker must have the means of knowing the IPs of the target MSs. Moreover, she must have means of knowing which MSs are in Sleep/Idle mode. If she chooses devices randomly then only a part of them will be in Sleep/Idle mode and be able to harm them. Also, according to the specification MSs waking up from Sleep mode typically do not conduct initial ranging as the authors claim [28]. As a matter of fact the MS performs the procedure for periodic ranging on regular basis even during Sleep mode. Another point of concern is that, the specification indicates that when there is traffic available for a specific MS the BS can choose between sending a MOB_TRF-IND immediately or buffer the data and notify the MS at a later time. This means that when the network is under heavy load the BS may choose to postpone waking up some of the devices. For the reasons mentioned above this attack should by considered unattainable for the WiMAX realm.

2) MOB_TRF-IND Water Torture Attack: This attack takes advantage of the unauthenticated MOB_TRF-IND normally sent from a BS to a sleeping MS when there is traffic pending for that MS. Skipping sleep mode is expected to reduce the battery life of the MS. This means that, activating and deactivating device components when falling to sleep and waking up from it has an energy cost of its own. If the attacker is able to forge a valid MOB_TRF-IND, and repeatedly transmit it to a sleeping MS in the vicinity she would be able to drain the energy resources of the victim on a higher pace. This attack was first described in [15] and is also mentioned in [25], [29].

The implementation cost of this attack is high as an equipment acting like a BS is a necessity. Of course, this attack is possible for a limited number of MSs (those that have fallen into sleep mode) within the action range of the attacker's equipment. Since a complete or significant energy drain requires a lot of time, the attacker is running the risk of getting physically exposed while there is also a scepticism about the amount of energy (and therefore the level of annoyance caused to the victim) that can be consumed in this way. That is why this attack is further evaluated in section V. Therefore, this attack is classified as minor.

3) BR and UL sleep control header DoS Attack: As already explained, it is possible for an MS to request activation of a sleep mode by sending a BR and UL sleep control header instead of the more common way with a MOB_SLP-REQ message. The authors in [30] claim that it is feasible for an attacker to forge a BR and UL control header with the victim's identity (MAC Address) and send it to have that MS fall into sleep mode. As a result, the BS will stop transmitting messages to that MS and DoS will occur.

In our opinion, this is actually an implementation-dependent feature. The specification indeed leaves room for such an invalid request for sleep to happen even though in careful implementations the BS is expected to reject or postpone any requests for Sleep mode if the BS has currently queued traffic pending for that MS. In the best case scenario for the attacker this will be expressed as some disturbance for the user that will include a brief lack of service for the MS to fall into sleep and the wake up after the first availability window. Moreover, the attacker must know whether the victim's equipment does support sleep functionality. At this point we have to concede that the authors in [30] managed to reproduce this attack on a 802.11 network and observed that its results was lack of connectivity for the victim MS. The IEEE 802.11 and 802.16 standards indeed share very similar Sleep Mode mechanisms but we believe that as already mentioned this is a vendor specific issue and that at least for the 802.16 case, if the implementation is done by the standard then there will be no room for substantial DoS. Taking into account all the above, this attack should be considered as a minor one.

4) Secure LU DDoS Attack: Location Update (LU) is a process by which a BS stays informed about the current location of a given MS. This process may be initiated by the MS at will or when one of the following conditions apply: (a) the MS detects a change in paging group, (b) periodically, prior to the expiration of the idle mode timer, (c) as part of its power down procedure, and (d) when the MS MAC hash skip counter exceeds a threshold. There are two modes supported: secure LU or unsecure LU. In secure LU the MS is required to send an RNG-REQ message to the BS including a HMAC/CMAC tuple. As always, the BS will have to verify the HMAC/CMAC value. If the current BS does not share security context with the MS then it will request it from the backbone network via the LU Request message. The backbone network will generate and provide the keying material via an LU Reply message.

The authors in [31] claim that this process can strain the network when it is performed simultaneously by a large number of devices. Since any MS can request bandwidth for LU, the attacker will simply have to construct a valid RNG-REQ message but with wrong HMAC/CMAC. A malicious MS can generate a large number of requests easily and without running the risk of getting discovered.

In principle, this attack is very similar to the RNG-REQ DDoS one but it involves some additional procedures (named above) by both the BS and the backbone network that may magnify the result and cause additional damage. Taking that into account, this attack can be registered as major.

C. Handover Attacks

Hand Over (HO) is the process in which an MS is transferred from its current BS to the air-interface of another neighboring BS. HO is a multi-step process comprised of the following main steps:

- Cell reselection - An MS evaluates neighboring BSs as candidates. To do so, the BS transmits a MOB_NBR-ADV message periodically that contains the appropriate information. This allows an MS that seeks handover to identify all the BSs in the neighbourhood.
- HO Initiation - An HO process may initiate either by the MS or the serving BS. In the first case, the intention is declared with a MOB_MSHO-REQ while in the second case by a MOB_BSHO-REQ one. Note that the handover command message includes one or more target BSs.
- Synchronization to new BS - The MS will synchronize to the DL of the target BS and gain its DL and UL parameters.
- Ranging - Either full initial or HO ranging may be conducted between the MS and the target BS. Depending on the information the target BS already has about the MS it may decide to skip one or several steps of the ranging process.
- Termination of MS Context - The serving BS (old) shall terminate all connections addressed to the MS and all contexts associated with them (i.e., information in queues, ARQ state machine, counters, timers, header suppression information).

1) MOB_NBR-ADV Downgrading Attack: Since this type of messages is not integrity protected, the attacker is able to alter them by removing information about neighbour BS in the appropriate message fields. This will prevent the handover procedure to happen as the victim MS will think it is isolated. While moving away from the serving BS the MS will have no other choice than to remain attached to it and the QoS will be reduced gradually until it will be out of service [15], [24], [17], [25].

A BS like equipment is necessary to the attacker thus increasing the implementation cost. The attacker must have preestablished a tunnel between the MS and the BS, constantly eavesdropping for any MOB_NBR-ADV messages and then transmitting simultaneously but with stronger signal so that her message be processed and not the valid one. It is easy to realise that the attack focuses mostly on single target MSs. Since this message is transmitted periodically and the MS is expected to move, the attacker must also follow the MS movement and always alter these messages when they are sent. The implementation methodology is rather difficult and is expected to demotivate attackers which are simply aiming at causing disturbance to a specific user. Therefore, the MOB_NBR-ADV Downgrading Attack should be considered minor.

2) MOB_NBR-ADV DoS Attack: Another option for an attacker is to manipulate the MOB_NBR-ADV in a way that will announce the presence of a non-existing BS with better characteristics than the serving one [15], [17]. The authors claim that this will cause DoS for the legitimate users as the MS will disconnect from its currently serving BS when trying to connect to the new one that does not exist. In a more dangerous case the attacker will include the information of rogue BS she controls thus possibly associating a legitimate user with it.

As a fact the specification does support soft HO. According to it the termination of the connection with the serving BS happens only as a final step of the HO and only after the MS has synchronized with the new BS. If that does not happen the MS will simply not abandon its current BS. Soft HO is more appropriate (and therefore expected to be enabled in high mobility networks). Nevertheless, the soft HO mode is not the default one. Unlike soft HO the hard HO mode leaves the field open for this attack to cause persistent DoS to specific MSs. The real danger lies in the fact that it gives an attacker the potential to associate with a malicious BS and from there launch more severe attacks. Overall, by itself the attack is only a minor one.

D. Miscellaneous Control Message Attacks

1) SBC-REQ Security Downgrade Attack: As seen in section II-B, during the initial network entry basic capability negotiation takes place. During this step among other parameters the MS informs the BS about the supported security capabilities of the device. This is carried out via a negotiation process that involves the exchange of SBC-REQ message from the MS to BS and the SBC-RSP message from the BS to the MS. The attributes that may be included in these messages are:

1) *PKM Version Support* - Indicates the supported versions of the PKM protocols. PKMv1 corresponds to the first bit of the field, while PKMv2 corresponds to the second bit. For each bit a value of 1 indicates that the corresponding protocol is supported while a value of 0 not supported.
2) *Authorization policy support* - Indicates the authorization policy to be supported by the MS and BS. This can be EAP only authorization, RSA only authorization, EAP authorization after RSA authorization. This is set by the value of the first two bits of the message in combination. In the SBC-REQ message this attribute is used to inform the BS about all authorization policies the MS can support but only the BS can decide and set the active one in the subsequent SBC-RSP. If all bits of this attribute included in the SBC-RSP message are zeroed, then no authorization may be applied but this is totally up to the BS.
3) *MAC mode* - Indicates the MAC modes the MS supports. As already mentioned, the possible values are HMAC (64-bit, 80-bit and 96-bit) and CMAC. The MS should support at least one message authentication code mode. It is possible that all bits of this attribute are 0 (this is possible only for the SBC-RSP message produced by the BS). In this case no message authentication code may be applied.
4) *PN window size* - An MS or BS maintains a record of the PN of the latest PDU received for each SA. This is done to avoid replay attacks. PDUs with a PN smaller than the beginning of a PN window shall be discarded as a replay attempt. So, this attribute defines the window size.

5) *PKM flow control* - Indicates the maximum number of simultaneously active PKM transactions. The value can be from 1 to 255 with default value of no limit (0 value).

Since these messages are exchanged before the BS and MS start an encrypted session, no actions for securing the contents of SBC-REQ exist. This vulnerability was first mentioned by [32] and later on in [16]. Also, the authors in [15] described a potential attack by exploiting this vulnerability. An attacker may attempt to alter a valid SBC-REQ message send from a legitimate MS to the serving BS during the network entry process. The forged message should contain false information about the security capabilities of the legitimate SS, typically lower or no security capabilities. The authors claim that in the second case, the communication between the two parties will be conducted in a non-encrypted way, allowing any malicious entity to easily eavesdrop the communication.

Although this message is indeed unencrypted and unauthenticated (which makes message spoofing possible), we argue that a security level downgrade attack with results comparable to what the authors describe, is highly unlikely to occur in this way. It is true that the specification does leave open the possibility for a session to be conducted without authorization. This possibility exists mainly as a safety net for situations that a BS is facing technical problems at a given moment and is incapable of supporting such procedures (e.g. because the backbone network faces a problem or due to a database failure) or for letting the BS operating under emergency mode. Ultimately, this is something only the BS decides and enforces through the SBC-RSP message. In realistic conditions it is not likely that a service provider will allow an MS to register to its network without supporting any authorization policy or any MAC mode. Thus, the attacker has little hope of succeeding by simply setting the Authorization Policy Support and MAC Mode fields to zero. Besides that, during the three-way handshake an MS sends the SA-TEK-Response message which includes the Security_Neg field. The message is encrypted and authenticated. Upon receiving this message the BS checks if the security capabilities contained in the SBC-REQ match the ones contained in SA-TEK-Response. If not, the protocol will report this inconsistency to higher layers, thus the BS will be aware of an attempted attack. This protection mechanism exist only as part of the PKMv2 authentication (PKMv1 does not have a similar verification process). Thus an attacker is bounded necessarily to at least change the field that indicates the PKM protocol version if she wishes to go undetected. Actually, downgrading to PKMv1 is a sure bet for an attacker. Although by doing so the attacker may never achieve non-encrypted communication, PKMv1 is known for many vulnerabilities, thus this practice could be used as the first step to launch more serious attacks. Overall, this attack should be considered infeasible or minor.

2) FPC Downgrade Attack: Fast Power Control (FPC) is a mechanism used for simultaneously adjusting the power levels of many MS to an optimal level. In this perspective it resembles periodic ranging although it is much faster. FPC messages are sent on the Broadcast CID. The format of this message is rather simple. It contains the number of MSs to be affected and for each MS its Basic CID and the power correction.

This management message is not integrity protected thus it may be altered to set the transmission power of the MS to non optimal levels, either too low or too high. In the first case, the MS will go through the procedure of adjusting its power levels until the signal is strong enough. The simultaneous power adjustment messages will result in many uplink bandwidth requests. This generally causes collisions in uplink of the MS and stalls the procedure of acquiring correct transmission power. This case resembles the RNG-REQ Downgrading Attack [25], [29].

If the BS supports FPC then indeed the aggressor by simply broadcasting a single message in specific time intervals may be in position to create disturbance or even DoS to moderate number of MSs within her action range. This implies that the attacker has invested funds in acquiring a strong BS like equipment for transmitting these messages. By having multiple attackers collaborating this attack can be easily extended to affect more users and eventually transforming it into one that could overstress the BS. Nevertheless, the implementation of the FPC is optional for the BS which reduces significantly the motivation of the attacker to launch it against any given network. It is unlikely this attack should be considered major.

3) FPC Water Torture Attack: This is a slightly modified version of the FPC Downgrade Attack. This may lead to drain the batteries of MSs [17], [33], [24], [25] using a methodology and characteristics much similar to the ones of RNG-RSP Water Torture Attack. Hence this attack should also be considered minor.

4) RES-CMD DoS Attack: Reset Command (RES-CMD) is a message that is used as a mechanism to reset an MS that appears to be unresponsive to its serving BS, or there are persistent anomalies in the UL transmission. When this message is received by an MS then it shall reset itself, reinitialize its MAC, and repeat initial network entry procedure.

This message is protected by HMAC/CMAC therefore it cannot be spoofed by an attacker and be sent at any time. Nevertheless, it is possible to force the BS to transmit this message by itself. It is easy for an attacker to learn the burst profile of a specific MS, through the UL-Map message. Upon that she can systematically choose to transmit at the exact same times as the victim SS. Provided that the two signals will arrive at the BS with similar power strength the final message the BS receives will appear as a single unintelligible message. The BS may be fooled into considering that the MS is malfunctioning and therefore issue a RES-CMD command [24].

In this way, it is possible for an attacker to achieve disturbance or even DoS against a single target MS. Since the result of the RES-CMD command generally triggers procedures that involve heavy signalling operations, an attacker might be tempted to expand this attack to larger scales. Due to the fact that the attacker must monitor and act in very tight timeframes in coordination with its victim, such implementation might prove quite hard in real life as it assumes an one-on-one victim-attacker ratio. Realistically, the challenging implementation methodology may discourage an attacker from adopting such approach for larger number of users. Thus, this attack is classified as one of minor level.

5) DBPC-REQ DoS Attack: Downlink Burst Profile Change Request (DBPC-REQ) is a message transmitted from the MS to the BS on the Basic CID to request a change of the least effective DL burst profile. Usually, this happens when channel conditions change. Actually, the DBPC-REQ message itself can be used as a quicker alternative to the RNG-REQ message [25].

In fact DBPC-REQ message is also unauthenticated. An attacker can change the Burst profile (modulation, encoding etc.) with the intention of disrupting communication between the BS and MS by the misuse of the DBPC-REQ message.

We can argue that this attack is similar to the RNG-REQ Downgrading Attack in matters of methodology and possible effects and for the same reasons should be classified as minor.

E. Attacks Against WiMAX Security Mechanisms

1) Interleaving Attack: This attack was mentioned by [34], [35] and consists of two rounds. In the first round the attacker impersonates a valid MS and sends an Authentication Information message followed by an Authorization Request message which have been intercepted and stored from a previous valid session of that MS. After receiving the Authorization Reply message the attacker must complete the authorization protocol by providing a valid Authorization Acknowledgement response. The attacker is not in position to construct this message because she does not have knowledge of the valid MS's private key and cannot decrypt the Authorization Reply message. However, the attacker can start the second round (in parallel with the first round) aiming at using the valid MS as an oracle to construct an Authorization Acknowledgement message on her behalf. In this round the attacker will take the role of a BS. By forcing the MS to start another protocol instance, it will use the Authorization Reply of the first round (it was received by the valid BS). The valid MS will provide the correct Authorization Acknowledgement message which the attacker will forward to the valid BS and finish the first round.

Protocol 4 Interleaving Attack

EVE $\rightarrow$ BS: ($Cert_{Manufacturer}$)
EVE $\rightarrow$ BS: (Rnd_{SS}, $Cert_{MS}$, [Cryptographic Capabilities], Basic CID, $Signature_{SS}$)
BS $\rightarrow$ EVE: ($Cert_{BS}$, $Enc(pre-PAK)_{PK_{SS}}$, Sequence Number, PAK Lifetime, [SAIDs],Rnd_{SS}, Rnd_{BS}, $Signature_{BS}$)

MS $\rightarrow$ EVE: (Rnd_{SS}, $Cert_{MS}$, [Cryptographic Capabilities], Basic CID, $Signature_{SS}$)
EVE $\rightarrow$ MS: ($Cert_{EVE}$, $Enc(pre-PAK)_{PK_{SS}}$, Sequence Number, PAK Lifetime, [SAIDs],Rnd_{SS}, Rnd_{BS}, $Signature_{BS}$)
MS $\rightarrow$ EVE: (Rnd_{BS}, Authentication Result, Error Code, Display String, $Signature_{SS}$)

EVE $\rightarrow$ BS: (Rnd_{BS}, Authentication Result, Error Code, Display String, $Signature_{SS}$)

In this way the attacker having acted as a simple Man-in-the-Middle entity will authenticate herself rather than the valid SS and trick the system into registering the wrong user. Nevertheless, no real gains in terms of theft of service can be achieved in this way. The attacker will still not be in possession of the AK, TEK or any other keying material and therefore will not be able to decrypt traffic sent by the BS or construct messages with valid HMAC/CMAC. In the best case, she can only continue to act as a Man-in-the-Middle and manipulate the valid MS-BS conversation by dropping or even forging unprotected control messages more easily. We argue that in most of the cases there is no motivation for an attacker to launch this attack.

2) AUTH-REQ Replay Theft of Service Attack: The authors in [17] claim that the random number field contained in message Auth-Req, fails to protect against reply attacks. The message can still be retransmitted by an attacker and the BS will have no means of knowing about its freshness.

Actually, the random number field in the Auth-Req message is a mechanism introduced to associate each Auth-Rep message with one Auth-Req and not to protect Auth-Req from replay attacks. The MS will know for sure that the Auth-Rep is fresh, if the MS random number field matches the one originally sent in the Auth-Req message. We argue that even though it is possible for an attacker to replay this message, there is no real gain involved for her. Indeed, an attacker may send this message in an attempt to authenticate herself. The message will be well formed and the BS will not be able to judge if it has been sent in the past by another SS. However, the subsequent message carrying the pre-PAK can only be decrypted with the valid SS's private key which is not revealed. Therefore, there is no real threat associated with this attack.

3) AUTH-REQ Replay DoS Attack: Xu and Huang [36] presented this attack against the first version of the PKM protocol. In this attack, the attacker stores and replays an instance of the Auth-Req message a legitimate SS has sent in the past. It is possible that a BS has set a timer that forces it to reject duplicate Auth-REQs originating from the same SS within a specific period. This means that the BS might as well drop legitimate requests coming by the victim SS. Depending on the vendor it is possible for this attack to be feasible in the PKMv2 of the protocol. In this case there are two possibilities: (a) either the attack will take the course the authors described leading to a DoS against a small/moderate number of users, or (b) the BS will proceed normally with the authorization process giving room to collaborating attackers for DDoS attack. This possibility was also recognized in [37].

Considering the second case, for each Auth-Req message the BS will have to verify each of the messages signature, generate keying context, construct the Auth-Reply message and finally transmit it to the MS. It is obvious that this sequence of actions can be a burden to the BS especially if it is repeated many times or for a large number of simultaneous requests. The problem with this attack which differentiates it from other DDoS attacks against WiMAX, is that it has an upper limit. By this we mean that there is a limit on the number of simultaneous requests that collaborating attackers may issue. This is mostly due to the fact that the Auth-Req message contains the SAID field which will be checked and used for the construction of the Auth-Rsp. This practically limits the attacker to replay Auth-Req messages of only those MSs whose basic CID is still active. Theoretically, the challenge of this attack is to have N number of collaborating attackers issuing simultaneously M number of requests where M is significantly smaller than the number of simultaneous connections a BS can support. It is very interesting to see if such number of requests requires enough computational

resources that the BS would stop responding for a period of time. The relatively small possible gains retrieved from our experimental results allow us to classify it as moderate (see section VI).

4) PKM-RSP: Auth-Invalid DoS Attack: PKM-RSP are messages issued by the BS and sent to the SS. Generally, messages of this kind are comprised by the following fields:

- Management Message Type - for PKM-RSP messages the value of this field is 10.
- Code - this field identifies the type of PKM packet.
- PKM Identifier
- TLV Encoded Attributes - PKM attributes carry the specific authentication, authorization, and key management data exchanged between the MS and BS.

The Auth-invalid message is sent by BS to MS when (a) the AK shared between BS and MS expires or (b) the BS is unable to verify the HMAC/CMAC properly. This message has a great chance to be used as a DoS tool for shutting out legitimate subscribers. First of all, the Auth-invalid message itself is not HMAC/CMAC protected. Also, it is sent unsolicited from the BS (when one of the aforementioned conditions occur). The PKM identifier mechanism -generally used for correlating response messages to the appropriate requests thus protecting from fake ones- is not used in this case. Thus, the attacker can easily inject fake messages of this kind into the channel and be certain that an MS will not recognize them as such. If the MS accepts the message it will pass to Reauth Wait state waiting further instructions from the BS. In case where the MS's Reauth Wait timer expires without receiving a message from the BS then the MS will send a Reauth Request in order to restart the authentication process. In the Reauth Wait state the device may accept messages such as an Auth Reject which will cause immediate break of all subscriber traffic [17], [24], [38]. Having in mind all the above, this attack is characterized as major due to its simplicity and immediate effect.

5) TEK Reuse Attack: This vulnerability was introduced in [39] and was supported by a large number of papers later on [35], [40], [41], [37]. Most researchers recognise that the TEK identifier (which its length of 2 bits is extremely short) can be used to identify only four generations of TEK keys. It is stated that due to this fact it is straightforward for an attacker to inject or replay expired TEKs.

Unfortunately, an analytical attack methodology is never appended in any of the aforementioned works. Moreover, to the best of our efforts we could not identify any means to exploit this vulnerability. Both messages responsible for the delivery of the TEK to the MS have it encrypted by the KEK. These messages are also protected by HMAC/CMAC and additionally they are secured by replay protection mechanisms such as random number challenge-response. Therefore, we believe that there is no real danger imposed to the system by performing this attack.

6) DES CBC IV Attack: Cipher-Clock Chaining (CBC) is a cipher mode of operation in which the plaintext is broken into a number of fixed size blocks and each one of them is XORed with the previous ciphertext block before being encrypted. In this way, each message is unique and each block is dependent on all preceding plaintext blocks. The first block is a special case where no previous ciphertext block exists, thus an IV must be used. More specifically given a symmetric key K and a plaintext P:

$$C_i = E_K(P_i \oplus C_{i-1}), C_0 = IV$$
$$P_i = D_K(C_i \oplus C_{i-1}), C_0 = IV$$

Generally, it is important for the IV to be unique and unpredictable. If not unique, then the CBC mode is degraded to a simple Electronic Codebook (ECB) mode where the distribution of the sequences of characters is maintained thus allowing traditional cryptanalysis methods (such as statistical analysis) to succeed. If not unpredictable, then it gives room to a chosen plaintext attack to succeed.

It is true that while the IV for AES in CBC is produced in a secure way the same is not true for the IV that is used for DES in CBC mode. Actually, as described in section II-C5 the specification states that the IV field in the keying information should be generated in a random way (we can safely assume as random as the rest of the keying material) and then be XORed with the frame number or the UL-MAP for the DL and UL respectively. The IV field is static for the entire TEK lifetime and is transmitted as an unencrypted field of the RSP: Key Reply message. On the other hand, the frame number is a simple counter and this makes the final IV material predictable. This vulnerability was first discussed in [39]. Although an attack methodology was never given we assume that the author implies a known plaintext attack. In attacks of this type the aggressor typically, as a first step, captures a cipher block c_v. Next, she generates a plaintext block of information $p = IV_i \oplus IV_{i+1} \oplus P_{guess}$ where IV_i the IV used to construct the c_v, IV_{i+1} is an estimation of what the next IV is going to be and P_{guess} is a guess of the plaintext encrypted to produce c_v. Then the attacker sends and forces the victim to encrypt p as follows $c_a = E_K(IV_i \oplus IV_{i+1} \oplus P_{guess} \oplus IV_{i+1})$ and as a last step she compares the two ciphertexts. If $c_v = c_a$ then her original assumption about the plaintext block will be true. It is to be noted that, similar attack methodologies have been investigated for the IPsec realm in the past [42], [43].

We can conclude that this vulnerability leaves room for attacks which are mostly targeting the verification of an assumption of a plaintext rather than actually revealing unknown plaintext or breaking the encryption key. In most situations, this may proved to be of little value for an attacker. Moreover, there is high implementation complexity associated with the methodology of this attack as the attacker will face the issue of forcing the system to encrypt and transmit the chosen sequence of data. Thus, this attack is classified as minor.

7) DES CBC Insecurity Attack: This vulnerability was first revealed in [39]. According to the authors DES in CBC mode looses its security after $2^{n/2}$ blocks encrypted with the same key, where n is the size of the blocks used by the block cipher. Since DES utilizes 64-bit blocks, it is expected that after 2^{32} 64-bit blocks the security of the system will be diminished. This becomes feasible because, under realistic conditions WiMAX networks have data rates that exceed this security threshold before the end of the TEK's lifetime.

Although an analytical methodology for this attack is never provided in the literature, it is certain that the first step for the aggressor is to force the system to switch to PKMv1 and then

instruct it to choose DES in CBC mode. This is necessary because in PKMv2 the Authorization Request (which is the message that informs the BS for the supported cipher suites) is protected by the signature of the MS and cannot be forged (unlike the PKMv1 version of the same message which is not). The attacker must first send a bogus SBC-REQ message using the methodology described in section III-D1 and then transmit a fake Authorization Request message with the data encryption algorithm identifier field set to $0x01$. As a matter of fact, DES has been the center of extensive cryptanalysis and there are numerous works which point out weaknesses of the algorithm in theoretical level [44], [45]. Nevertheless, at least for the time being, most of the derived attack methodologies have unrealistic requirements (in matters of known or chosen plaintext size) and they have been proved impractical in real life situations. Still, the traditional brute force attack seems to be the most practical attack against DES. Since a direct citation with a proof of the statement above is never provided in any of these works, we can assume that the authors do not refer to any of the afformentioned attacks rather they imply the documented vulnerability of the DES-CBC cipher under birthday attacks. According to the birthday paradox, after $2^{m/2}$ block of data there is high probability that a collision will occur, i.e. two of the n ciphertexts will be the same [46]. A more detailed attack methodology of this type can be found in [47].

Given knowledge to the entire set of ciphertexts $C_1, ..., C_n$, if $C_i = C_j$ the two colliding ciphertexts, then the attacker can compute:

$$E_K(P_i \oplus C_{i-1}) = E_K(P_j \oplus C_{j-1}) \Rightarrow$$
$$P_i \oplus C_{i-1} = P_j \oplus C_{j-1} \Rightarrow$$
$$P_i \oplus P_j = C_{i-1} \oplus C_{j-1} = a$$

Where P_i, P_j are the corresponding plaintexts and a is a block of bits. Since the block size is 64 bits it is true that there is high probability (but not certainty) that some information leakage will happen after 2^{32} blocks of data. As stated in [46] this kind of information is non-trivial. On the downside with the knowledge of only $P_i \oplus P_j$, a statistical analysis over a packet as small as 64 bits is not practical. Even in the case where random a block of 64 bit in a set of 2^{32} blocks of information will be of little value. What is more, 2^{32} blocks of 64 bit long blocks equals to an amount of traffic that exceeds 34 GB. This size of traffic may indeed be exchanged in 12 hours with an average throughput of 6 Mbps but it is unlikely that any given client would be willing to exchange so large traffic in such a small timeframe. Thus, we believe that this attack is strictly theoretic and infeasible in practice.

F. Multicast/Broadcast Attacks

Multicast and Broadcast Services (MBS) was a major introduction of the 802.16e specification. MBS is an operational mode that allows the dissemination of data across multiple MS of the network, from a single centralized media server. Communication in MBS is unidirectional and is offered in the DL only. This practically means that a BS can send a message simultaneously to all the members of the same group. For transmission of encrypted content, it is obvious that keys commonly shared to all group members are necessary. In MBS these keys are known as Group KEK (GKEK). Provided that an MS is already registered and authenticated to the network, the BS will randomly generate and send the GKEK (encrypted with the KEK). The GKEK is used to encrypt the GTEKs sent by the BS to the MS that belong in the same multicast/broadcast group. Since GTEKs are known to many nodes, their proper maintenance and frequent refreshing is of paramount importance. MBRA is the mechanism employed for this purpose. An MS will request its GTEK from the BS by issuing a PKMv2 Key Request and will acquire this key through a PKMv2 Key Reply message. This pair of messages are carried on the Primary Management connection. A BS may request to update and distribute keying material in predefined time intervals by transmitting two PKMv2 Group Key Update Command messages to all group members. Two types of the PKMv2 Group Key Update Command message exist: the GKEK Update Mode and GTEK Update Mode. These two messages include a counter field, namely Key Push Counter, for protection against replay attacks. Figure 6 illustrates this procedure.

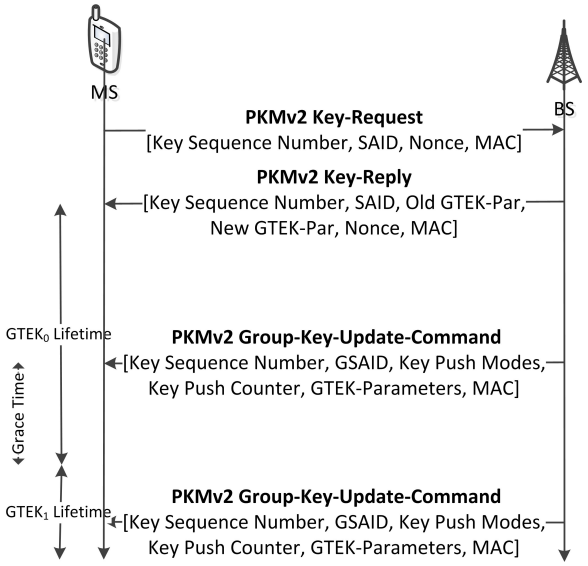


Fig. 6. MBRA messages

1) Group Key Update Command: GTEK Update Mode DoS Attack: GTEK is shared among all members of a multicast/broadcast group so that each member is able to decrypt the traffic it receives from the BS. GTEK is a symmetric key. This means that an MS cannot only decrypt data but also encrypt them using the same GTEK key. The members of the same group will be able to decrypt such messages but will not be able to distinguish if the message originates from the BS or an ill-motivated member of the group. Since this message has a valid encryption and HMAC/CMAC the other MS will take for granted that the traffic is originated from the legitimate BS. An adversary MS, member of the group, can use this opportunity to send malicious traffic pretending to be the BS.

Another more harmful attack insident appears when the same situation happens but this time with the GTEKs. The GTEK is encrypted and transmitted to all group members using GKEK, which is also known to all group members. An

TABLE I
EVALUATION AND CATEGORIZATION OF WIMAX ATTACKS.

Category	Attack	Threat	Cost	Difficulty	Risk	Duration	Size	Outcome	Protocol
Ranging Attacks	RNG-RSP DoS Attack	Major	Expensive	Easy	High	Long	Large	DoS	802.16-2009
	RNG-RSP Downgrading Attack	Minor	Expensive	Solvable	High	Long	Medium	Annoyance	802.16-2009
	RNG-RSP Water Torture Attack	Minor	Expensive	Easy	Moderate	Long	Medium	Annoyance	802.16-2009
	RNG-REQ Downgrading Attack	Minor	Expensive	Hard	Moderate	Long	Medium	Annoyance	802.16-2009
	RNG-REQ DDoS Attack	Major	Inexpensive	Easy	Low	Long	Large	DoS	802.16-2009
	MOB_ASC-REP DoS Attack	Minor	Expensive	Solvable	Moderate	Long	Medium	Annoyance	802.16-2009
Power Saving Attacks	Signaling DoS Attack	-	-	-	-	-	-	-	-
	MOB_TRF-IND Water Torture Attack	Minor	Expensive	Easy	High	Short	Medium	Annoyance	802.16-2009
	BR and UL Sleep DoS Attack	Minor	Inexpensive	Hard	Low	Short	Large	Annoyance	802.16-2009
	Secure LU DDoS Attack	Major	Inexpensive	Easy	Low	Long	Large	DoS	802.16-2009
Handover Attacks	MOB_NBR-ADV Downgrading Attack	Minor	Expensive	Hard	High	Long	Small	Annoyance	802.16-2009
	MOB_NBR-ADV DoS Attack	Minor	Expensive	Hard	High	Long	Small	Annoyance	802.16-2009
Miscellaneous Control Message Attacks	SBC-REQ Security Downgrade Attack	Minor	Expensive	Hard	Low	Long	Small	Loss of Privacy	802.16-2004
	FPC Downgrade Attack	Moderate	Expensive	Solvable	Moderate	Long	Large	Annoyance	802.16-2009
	FPC Water Torture Attack	Minor	Expensive	Easy	Moderate	Long	Medium	Annoyance	802.16-2009
	RES-CMD DoS Attack	Minor	Expensive	Hard	Moderate	Short	Small	Annoyance	802.16-2009
	DBPC-REQ DoS Attack	Minor	Expensive	Hard	Moderate	Long	Medium	Annoyance	802.16-2009
Attacks Against WiMAX Security Mechanisms	Interleaving Attack	-	-	-	-	-	-	-	-
	AUTH-REQ Replay Theft of Service Attack	-	-	-	-	-	-	-	-
	AUTH-REQ Replay DoS Attack	Moderate	Manageable	Solvable	Low	Long	Medium	DoS	802.16-2009
	PKM-RSP: Auth-Invalid DoS Attack	Major	Expensive	Easy	Moderate	Long	Large	DoS	802.16-2009
	TEK Reuse Attack	-	-	-	-	-	-	-	-
	DES CBC IV Attack	Minor	Manageable	Hard	Low	Short	Small	Loss of Privacy	802.16-2009
	DES CBC Insecurity Attack	Minor	Manageable	Hard	Low	Short	Small	Loss of Privacy	802.16-2009
Multicast/Broadcast Attacks	Group Key Update Command: GTEK Update Mode DoS Attack	Moderate	Inexpensive	Easy	Low	Long	Medium	DoS	802.16j
	GTEK Theft of Service Attack	Major	Manageable	Easy	Low	Long	Medium	Theft of Service	802.16-2009
	MCA-REQ DoS Attack	Major	Expensive	Easy	Moderate	Long	Large	DoS	802.16-2009
Mesh Mode Attacks	Malicious Sponsor Node Attacks	Minor	Inexpensive	Hard	Low	Long	Large	Annoyance	802.16e
	PKM-REQ: Auth Request Replay Attack	Major	Inexpensive	Easy	Low	Long	Small	Theft of Service	802.16e
	PKM-RSP Replay Attack	Major	Expensive	Easy	Moderate	Long	Moderate	Loss of Privacy	802.16e
	OSS Distribution Attacks	Major	Inexpensive	Easy	Low	Long	Large	Theft of Service	802.16e
	PKM-REQ: Key Request DoS Attack	Major	Expensive	Easy	Moderate	Long	Large	Dos	802.16e

For more information on how these results have accrued please refer to appendix and the online resources of the manuscript [1].

adversary that is already member of the group can manipulate MBRA to distribute its own fake GTEKs using the GKEK she normally owns. The messages will again be valid, and all the members will eventually replace their current keys with the fake one. After that, all the group members except the attacker will no longer be able to decrypt incoming traffic from the original BS [48].

This attack is straightforward in implementation and can affect all MSs within the same MBS group with a single alter/broadcast of a message (which is typically a moderate number of MSs). Moreover, it persists for as long as the current GTEK remains active but for prolonged effect the attacker must actively and continuously alter/forge the Group Key Update Command with a fake key. The BS has no means of knowing that the MS of a given group have another (wrong) key. Taking into account all the above, this attack is classified as moderate.

2) GTEK Theft of Service Attack: New members joining a multicast/broadcast group are given the currently active GTEK. This means that besides being able to decrypt subsequent traffic, they are able to decrypt traffic sent in the past, beginning from the point when the specific GTEK became active. Therefore, an attacker can passively store traffic and near the end of the GTEK lifetime join the network as a valid user [48], [36], [49].

The methodology of this attack is extremely simple and does not require any costly equipment from the attacker's side. The actual service duration that the attacker will be able to intercept traffic is provider specific as the GTEK lifetime is not specified by the standard. Typical implementations set this counter anywhere from 30 minutes to 7 days which is adequate considering that MBS deals mostly with multimedia services. This attack highlights the issue of lack of backward secrecy of the MBRA. This attack is also considered as major.

3) MCA-REQ DoS Attack: Multicast Assignment Request (MCA-REQ) is a message sent from the BS to an MS requesting from it to join or leave a multicast polling group. Upon receiving this message the MS shall add the multicast CID to its transmission opportunities or remove it according to the Join/Remove command of the corresponding field. Subsequently, the MS will respond by sending an MCA-RSP message back to the BS. All these messages are transmitted over the primary management connection. Also, since the MCA-REQ message is sent unprotected an attacker may remove an MS from a polling group at will. This attack when executed against a single user, will be expressed as some disturbance, but DoS will not necessarily occur as the MS will use the mandatory contention based bandwidth allocation algorithm to request UL bandwidth. When this attack is done in larger scale then it is possible to cause overloading of the UL resulting in greater uplink delay [25]. Based on the previous analysis, the attack is considered as a major one.

G. Mesh Mode Attacks

Mesh mode is a special type of operation of a WiMAX network where traffic can be routed through other MS instead of the sole BS and MS communication. It is to be noted that although Mesh mode was introduced in 802.16-2004 and supported in the 802.16e-2005, from the 802.16-2009 version of the specification, support for Mesh mode ceased to exist. Therefore, Mesh mode specific attacks will be presented here, briefly, only for reasons of completeness. Works like [50], [51] explore such attacks in further detail.

Mesh mode also provisions for operations such as access control and authentication, thus a centralized node (BS) must exist to perform these tasks. If this node is not directly reachable then simply the node that requires authentication will use existing members of the network to reach it. When entering a mesh network a newcomer shall select one node that will act as intermediate with the BS and facilitate the process of authentication. This node is characterised as the Sponsor node and its role is to simply forward the authentication

messages between the new node and the BS in a transparent way. The authentication process will normally result in the new node acquiring the authorization key which in this case is characterized as Operator Shared Secret (OSS). This same key is shared among all members of the mesh network. As a second step the new node will attempt to exchange TEK with its neighbours and proceed to link establishment.

.

1) Malicious Sponsor Node Attacks: Other potential attacks target the authentication and key exchange protocol. This family of attacks is harder to occur by outsiders as a UDP tunnel is established between the Candidate node and the Authorization center thus preventing eavesdropping and tampering of the messages. Nevertheless, the protocol assumes that the Sponsor node is loyal. This is a strong assumption that creates a great risk, as the Sponsor node is responsible for forwarding the authentication messages between the Candidate and the Authorization center. A malicious Sponsor can easily intercept or alter specific fields of the PKM-REQ: Auth Request or PKM-REQ: Auth Response resulting to a number of different attacks such as:

- By altering the field of cryptographic capabilities in the PKM-REQ: Auth Request message an attacker can achieve a Security Rollback attack (i.e. select PKMv1 instead of PKMv2).
- By modifying or removing the SA information field of the PKM-RSO: Auth Reply message an attacker can cause a DoS attack.
- By distributing the OSS attackers can allow unauthorized entities to join the network.
- By altering the OSS an aggressor can prevent new legitimate nodes to enter the network successfully.
- By reducing the OSS lifetime field a node is forced to update the OSS more often which leads to draining its battery faster.
- By issuing a PKM-RSP: Auth Reject, a message that is not protected by HMAC/CMAC, towards the Candidate nodes.

2) PKM-REQ: Auth Request Replay Attack: One of the major vulnerabilities of the mesh mode authorization process is that it does not provide replay protection against authentication messages. Let's consider the following scenario: Node A is registered at two mesh domains D_a and D_b. The attacker is only registered in one of the domains, D_a. By manipulating the protocol (with methods already explained) an attacker can be the Sponsor for this node when it tries to enter the network and store the PKM-REQ: Auth Request message. In the process the attacker will attempt to join domain D_b pretending to be node A by simply replaying the PKM-REQ: Auth Request message. The BS of domain D_b is not in position to distinguish if the message comes from a different authorization process and will transmit a PKM-RSP: Auth Reply one, thus disclosing the OSS key.

3) PKM-RSP Replay Attack: Many attacks can also be caused by the fact that the authorization process is unilateral, meaning that the Authorization center authenticates the Candidate node but not the opposite. Attackers can forge or replay PKM-RSP messages, completely impersonating the Authentication center.

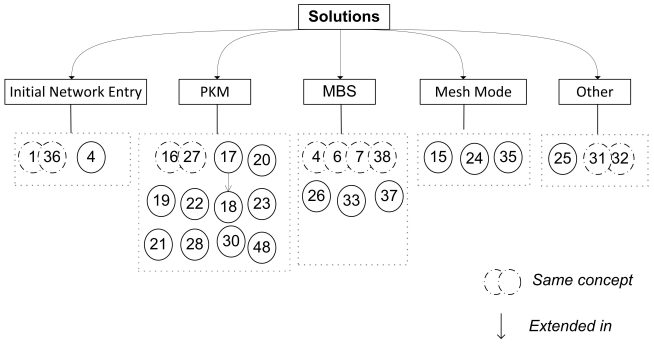


Fig. 7. Taxonomy of Solutions (Numbers Correspond to References)

4) OSS Distribution Attacks: Following the authorization phase, link establishment between nodes can take place using a challenge response protocol. In this protocol the cornerstone of confidentiality is the OSS key. Since this key is shared among all nodes in a mesh network, the assumption that all nodes will behave legally is too weak. There are no technical means that would prevent an attacker from disclosing the OSS to outsiders. With their turn these nodes can start establishing links with legitimate network nodes, without being authorized and request services from them.

5) PKM-REQ: Key Request DoS Attack: The fact that OSS is known to all network members can lead to even more attacks. Among others, OSS is used to produce the HMAC/C-MAC digests which guarantee the authenticity of several important messages. Recall that PKM-REQ: Key Request is a message used to distribute new TEK and the integrity of such messages is protected by HMAC/CMAC. A misbehaving member of the network that legally has possession of the OSS can forge PKM-REQ: Key Request messages with fake TEKs and use the valid OSS to apply the HMAC/CMAC. The victim node has no means to distinguish if the message really originates from the BS or any other node of the network. The result will be translated as DoS for the victim node as it will not be in position to decrypt the traffic.

IV. SOLUTIONS

This section organizes and analyses the solution proposed in literature to remedy some of the vulnerabilities of WiMAX. A taxonomy of the solutions that will be discussed is presented in figure 7.

A. Securing Initial Network Entry

Naseer et al. [17] focalize the cause of most DoS attacks in the unprotected management messages. The authors propose the utilization of Diffie-Hellman (DH) key exchange algorithm for the sensitive, unprotected process of initial network entry. Although, the DH algorithm is proven to be secure it relies in the common knowledge of a public key. The authors propose that parameters used for the generation of such public/private keys should be depended on Basic CID (BCID) and initial ranging codes. In this way, both the BS and MS can be sure of knowing the right chosen keys. Moreover, when no keying material is available, digital signatures should be used to authenticate management messages. However, this process

is vulnerable as it hides the risk for an attacker of knowing the BCID, ranging codes and then deducing the key pairs. A very similar approach is proposed by the authors of [38].

Han et al. propose a modified version of Diffie-Hellman protocol, namely Secure Initial Network Entry Protocol (SINEP), to secure the initial network entry procedure [15]. This protocol assumes that two variables (p a large prime number and r a primitive root of p) are shared a-priori by both the MS and BS. The authors claim that this challenge-response sequence when applied as an extra step to the network entry process it is able to secure against DoS and Man-in-the-Middle attacks. The protocol they propose can be formalized and presented in protocol 5:

Protocol 5 SINEP Protocol [15]

MS $\rightarrow$ BS: request
MS $\rightarrow$ BS: $H(H(ID_{SS}), nonce_{BS}, PK_{SS)}, PK_{SS}, nonce_{SS}$
BS $\rightarrow$ MS: $H(H(ID_{SS}), nonce_{SS}, PK_{BS)}, PK_{BS})$

B. Improving PKM

Xu and Huang propose an enhanced authentication phase of PKMv2 [34], [52]. This scheme makes use of timestamps in order to protect against simple replay attacks. Also, the proposed protocol has the advantage of involving less signalling overhead comparing to PKMv2. The authors also claim that synchronization is not a real issue since the BS and MS maintain synchronization during their ranging process.

Altaf et al. argue that timestamps impose additional security risks [35]. Actually, this risk has been studied in further extend in [53]. The authors present an authorization protocol of their own that is based on both timestamps and nonces. In this way security of the protocol does not depend solely on synchronized clocks. Sidharth and Sebastian [54] as well as Eren [37] agree that such hybrid solution may be a viable approach for an improved authorization protocol and propose their own similar versions. In the work of the latter it is also suggested that the supported size of the TEK sequence numbers should be increased from 2 bits ($2^2 = 4$ possible keys) to at least 12 bits ($2^{12} = 3,360$ possible keys). A formal representation of their protocol is given in protocol 6.

Protocol 6 Modified Authorization Protocol [35]

MS $\rightarrow$ BS: $Cert_{manufacturer}$
MS $\rightarrow$ BS: $Time_{SS}, Nonce_{SS}, Cert_{SS}, Capabilities_{SS}, BCID, Sign_{SS}$
BS $\rightarrow$ MS: $Time_{BS}, Nonce_{SS}, Nonce_{BS}, E(Pre-PAK, ID_{SS})_{PK_{SS}}$, $SequenceNumber, Lifetime, SAIDList, AAID, Cert_{BS}, Sign_{BS}$
MS $\rightarrow$ BS: $Time_{SS}, Nonce_{BS}, MAC_{SS}, Sign_{SS}$

Rahman and Kowsar propose a variation of the DH protocol to replace the existing authorization procedure in PKMv2 [55]. Their protocol assumes that each MS has a unique identity number and the BS has a different hash function associated for each MS. Additionally, global variables P and G necessary for the execution of the DH algorithm are required. Since the original DH protocol is vulnerable to Man-in-the-Middle attacks, the authors add extra steps on top of it to authenticate the parties. They also propose that the much simpler Vernam cipher [56] should be used instead of AES or DES, once both parties have acquired their common secret. Their scheme is illustrated in protocol 7.

Protocol 7 Modified Authorization Protocol Presented in [55]

MS $\rightarrow$ BS: Request
BS $\rightarrow$ MS: $Nonce_{BS}$
MS $\rightarrow$ BS: $H(Nonce_{BS}), ID_{SS}, Nonce_{SS}$
BS $\rightarrow$ MS: $H(Nonce_{BS})$
MS $\rightarrow$ BS: PK_{SS}
BS $\rightarrow$ SS: PK_{BS}

The possibility of utilizing Elliptic Curve Cryptography (ECC) for the authentication process of 802.16 had been studied earlier on. Fuqiang and Lei [57] propose a security framework based on PKI specifically adopted for the wireless realm. The proposed framework namely Wireless Public Key Infrastructure (WPKI) uses ECC instead of RSA in order to reduce the computation power substantially. A variation of the X.509 is also employed which is striped off any redundant information to save store memory. This authentication protocol may be lighter in matters of computational and memory resources comparing to the existing one but it does not deal with the attack described in III-E3. Generally several researchers have agreed that ECC might prove a viable security practise for the WiMAX networks [58].

Zaabi et al. propose an EAP-TLS-ECDH-RSA authentication mechanism (suite) which is a combination of user EAP-TLS authentication and device authentication based on ECDH-RSA [59]. As a first step, EAP runs but in order to acquire keys for building a TLS tunnel ECDH-RSA is executed. The BS creates an ECDH key but signs the corresponding certificate with RSA. The motivation behind this is the fact that RSA has proven to be slow in key generation while ECDH-ECDSA takes most of its time in verifying the certificate. Hence, the combination is expected to boosts the performance of the system. On the downside, the framework assumes that the MS will be able to generate an ECDH key on the same curve as the BS's public key. However, the MS may not always support such an ECC curve. This will result in a request on behalf of the MS, for the BS to renew its cryptographic parameters. In other words, the BS is required to renew its certificate. It is expected that more events of this type will have to be confronted as more MSs register with a BS. Works like [60] present similar authentication schemes based on the utilization of ECC.

Haibo et al. suggest that the key exchange phase of the PKM protocol should be modified, so that both the MS and BS contribute in the generation of the TEK [61]. Towards this direction, they propose a modified version of the Authenticated Key Exchange (AKE2) protocol [62] which achieves mutual control of the keys.

Li et al. present an alternative of the entire PKM protocol based on the concept of a Trusted Third Party (TTP) for its implementation [63]. In the first phase of their approach both the MS and the network (BS) need to register with a TTP server. The second phase involves MS and the network in exchanging their certificates and finally these two entities proceed with the exchange of the session keys. Also, the protocol assumes that both the MS and BS have acquired the TTP server's public key.

C. Enhancing MBS Security

From what is discussed in Section III it is clear that the MBRA is vulnerable to insider attacks where the owner of a shared GKEK can cause havoc to the rest of the group members. Moreover, it has been proven that the MBRA does not provide backward secrecy as any newly joining member can decrypt all traffic multicasted during the GTEK lifetime even if that member was not part of the group. Likewise, there is no provision for forward secrecy. This enables any member leaving the group, to continue decrypting traffic as well as be able to receive the next GKEK and decrypt the next GTEK. Another shortcoming of the MBRA which is not directly related with the security is its high complexity cost. Researchers proposing new protocols attempt to overhaul all flaws of MBRA while keeping its complexity cost as low as possible.

The authors in [17] point out that the security offered by the MBRA algorithm could be improved having the GTEK directly unicasted to each MS inside a given group. In this scheme, the GTEK is encrypted with KEK instead of GKEK, which is considered redundant. This scheme deals effectively with insider attacks since an attacker would require knowledge of the KEK. It also requires slightly less storage as less keys are saved. Obviously, this scheme maintains the problem of scaling poorly because N unicasts are required for a simple key refresh, where N is the number of members inside a group. Another major shortcoming is the lack of forward and backward secrecy. The protocol the authors propose can be formalized as in protocol 8.

Protocol 8 MBRA protocol proposed in [17]

Initial Keying: Not Considered
Key Update:
(1) BS $\rightarrow$ SS: $(GTEK)_{KEK}$
Rekeying at Join Event: Not Considered
Rekeying at Leave Event: Not Considered

Xu et al. adopt a similar approach to enhance the security of MBRA [49]. Their protocol also considers the existence of the GKEKs unnecessary and like the previous proposal it is based on unicasts of the GTEK. On initial key distribution N unicasts occur aiming to distribute the GTEK, plus one broadcast which serves as a simple notification. Unlike the proposal given in [17], here there is provision for join and leave events but in this case, key refresh is achieved with a single broadcast of a key update notification. This message triggers the MS to produce the new GTEK by itself by passing the old GTEK through a predetermined hash function. Normal key refresh happens with N unicasts. This protocol provides forward and backward secrecy, copes with insider attacks and requires roughly less key storage. In matters of efficiency it performs better comparing to the original MBRA although it is certain that for larger groups the performance will drop as it involves linear number of unicasts for a simple key refresh process. The proposed protocol is formalized and presented in protocol 9.

The work in [36] presents Elapse a modified version of MBRA based on the concept of subgrouping. According to this protocol, the existing MBS groups are further divided into a larger number of subgroups. The parameters of this

Protocol 9 MBRA protocol proposed in [49]

Initial Keying:
(1) BS $\rightarrow$ SS: $(GTEK)_{KEK}$
(2) BS $\Rightarrow$ SS: $Update\_Notice$
Key Update:
(1) BS $\rightarrow$ SS: $(GTEK)_{KEK}$
Rekeying at Join Event:
(1) BS $\Rightarrow$ SS: $Update\_Notice$
Rekeying at Leave Event
(1) BS $\Rightarrow$ SS: $Update\_Notice$

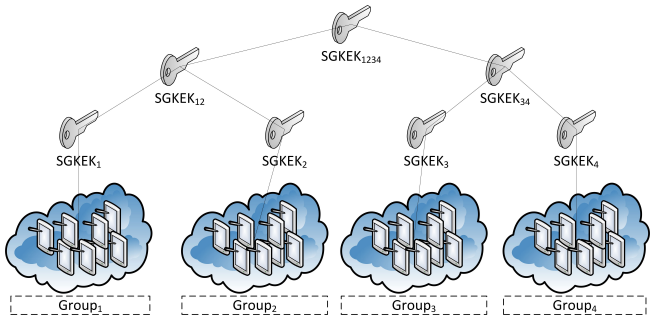


Fig. 8. Sample Subgroup Organization Including Key Hierarchy Proposed in [36]

organization including the actual number of subgroups and number of MS in each subgroup is manually decided by the administrator with respect to the individual requirements of a given application. All group members have the same GTEK but additionally, each MS belonging to a subgroup is also forced to maintain a chain of Sub Group Key Encryption Keys (SGKEKs) rather than a single GKEK. The total number of keys that must be maintained by an MS is increased to M (where M is the number of subgroups). A typical key refresh is covered by a single broadcast message rather than N unicasts plus 1 broadcast which is normally required by the current MBRA algorithm. Elapse provisions for the special cases of an MS joining and leaving a subgroup. In this way, perfect secrecy is provided. On the other hand, performance is expected to substantially decrease in cases where join or leave events happen on frequent basis within large groups. Therefore, very thorough and continuous administration of groups and subgroups is assumed. Another interesting point is that this protocol fails to resolve the insider attack issue described in section III-F1. A misbehaving member of the group is still able to encrypt traffic and distribute it acting like a BS since any member has knowledge of the GTEK. What is even more dangerous with the knowledge of the top level SGEK any group member is able to act as a BS and update the GTEK causing DoS until the next key refresh. The detailed sequence of messages according to the described protocol for an example of 4 subgroups is presented in protocol 10.

The rekeying algorithm proposed by Brown et al. [64] is based on a hash function scheme previously presented in [36] but adopted for the WiMAX realm. This scheme allows the MS to generate its keying material on their own, after key expiration rather than having the BS unicast or broadcast the corresponding keys. Although this proposal drastically improves the efficiency of the rekeying process by incorporating unicasts only during the initial keying phase, from a security point of view it should be considered very

Protocol 10 MBRA protocol proposed in [36] example for 4 subgroups

Initial Keying: Not Considered
Key Update:
(1) BS $\Rightarrow$ SSs: $(GTEK)_{SGKEK_{1234}}$
Rekeying at Join Event:
(1) BS $\rightarrow$ SS: $(SGKEK_{1234}, SGKEK_{12}, SGKEK_2)_{KEK}$
(2) BS $\Rightarrow$ SS_{SG3}: $(SGKEK_{1234})_{SGKEK_{34}}$
(3) BS $\Rightarrow$ SS_{SG4}: $(SGKEK_{1234})_{SGKEK_{34}}$
(4) BS $\Rightarrow$ SS_{SG1}: $(SGKEK_{1234}, SGKEK_{12})_{SGKEK_1}$
Rekeying at Leave Event
(1) BS $\rightarrow$ SS: $(SGKEK_{1234}, SGKEK_{12}, SGKEK_2, GTEK)_{KEK}$
(2) BS $\Rightarrow$ SS_{SG3}: $(SGKEK_{1234}, GTEK)_{SGKEK_{34}}$
(3) BS $\Rightarrow$ SS_{SG4}: $(SGKEK_{1234}, GTEK)_{SGKEK_{34}}$
(4) BS $\Rightarrow$ SS_{SG1}: $(SGKEK_{1234}, SGKEK_{12}, GTEK)_{SGKEK_1}$

weak. On the one hand, it suffers from the vulnerability against insider attacks as all group members share the same symmetric GTEK. On the other hand, it fails to provide forward and backward secrecy. This is true since the BS has to broadcast the structural elements of the new key, namely the random number. Although, this message is transmitted encrypted the encryption key (GTEK) is already know to a member that has just left the group, making it possible for her to extract the random number and deduce the new key. Besides that there are other indication of poor design. For example, the broadcast message during a leave event indicates that GKEK is common and shared among all group members. Since a key with the same characteristics already exists (GTEK), the GKEK is unnecessary. The formalized version of the scheme described above is given in protocol 11.

Protocol 11 MBRA protocol proposed in [64]

Initial Keying:
(1) BS $\rightarrow$ SS: $(GKEK, GTEK)_{KEK}$
Key Update:
No Transmition
Rekeying at Join Event:
(1) BS $\Rightarrow$ SS: Rnd_{GKEK}
Rekeying at Leave Event
(1) BS $\Rightarrow$ SS: Rnd_{GKEK}

The works in [33], [25] claim that broadcasting GTEK poses a vulnerability for the existing MBRA. The authors propose three possible modification of the algorithm to improve the security of MBRA. As a first option it is proposed that the Group Key Update Command: GTEK Update Mode message should be unicasted similarly to the Group Key Update Command: GKEK Update Mode message. This process should be performed automatically by the BS, i.e. without any request message sent by the SS. According to the authors this saves half of the bandwidth, comparing to the original MBRA algorithm. This approach is very similar to the ones described in [17] and [49] and carries the same advantages and inefficiencies. The second approach that is proposed involves the use of public key cryptography. The Group Key Update Command: GTEK Update Mode message is modified so that a digital signature of the BS is appended. In this way insider attacks can be avoided as the MS can verify that the author of the update messages is the legitimate BS. Unfortunately, forward and backward secrecy is not guaranteed and there is a slight additional computational cost due to the verification of asymmetric signatures. Similarly to the original MBRA this scheme scales poorly.

Finally, another solution that is proposed by the authors is the generation of GTEK as part of a hash chain. Initially, the BS shall have to generate GTEK from a random number. This key is considered the first generation GTEK and is represented as $GTEK_0$. Subsequent GTEK are created by applying a one-way hash function to the GTEK of the previous generation according to the equation $GTEK_n = f(GTEK_{n-1})$. Unlike the popular approaches of this kind, this scheme dictates that the last (i.e. the new) $GTEK_n$ is transmitted to the MS and not generated. The MS reproduces the $GTEK_n$ only for verification purposes and if the keys do not match the MS should employ the unicast Request/Reply mechanism. The introduction of the active GTEK verification only partially solves the problem of insider attacks. Indeed, a malicious MS will not be able to fabricate false keys and distribute them, but will be able to encrypt and distribute fake traffic. Actually, this mechanism leaves room for a new DoS attack to occur. Consider the following scenario: A misbehaving node inside the network decides to fabricate a Group Key Update Command: GTEK Update Mode message with a false GTEK. After broadcasting this message to all other nodes of the group each MS will reject it and initiate a unicast Request/Reply sequence roughly at the same time. The multiple unicasts will probably create a heavy signalling load which is possible to result in DoS. Moreover, this scheme provides only partial forward and backward secrecy since it enables a misbehaving node that has just joined/left the network, to decrypt all data since the last hash chain generation. Finally, it fails to reduce the high communication cost of the original MBRA. The messages of the second protocol are illustrated in protocol 12.

Protocol 12 The Second MBRA protocol proposed in [25]

Initial Keying: Not Considered
Key Update:
(1) BS $\rightarrow$ SS: $(GKEK)_{KEK}$
(2) BS $\Rightarrow$ SS: $(GTEK)_{GKEK}, Signature_{BS}$
Rekeying at Join Event: Not Considered
Rekeying at Leave Event: Not Considered

Kambourakis et al. [48] propose an improved version of the MBRA (Multicast/Broadcast Rekeying Algorithm) based on asymmetric encryption methods, such as ECC and bilinear maps. Instead of distributing a shared secret group key the protocol assumes that each member has a publicly known key (available even to attackers). The protocol results on each member acquiring a secret decryption key (unique to each member) rendering only the valid nodes able to decrypt these messages. The encryption practices of this protocol in principle are quite similar to the ones of the ElGamal algorithm. The proposed protocol deals effectively with insider attacks and provides both backward and forward secrecy. In matters of efficiency it requires more computational power due to the extensive use of asymmetric cryptography although, the computational cost is kept on reasonable levels because of ECC. In matters of communication overhead it performs well in join and leave events since it requires only one broadcast on behalf of the BS but scales poorly in normal key updates since it requires N unicasts on behalf of the MS plus 2 broadcasts from the BS. The fact that asymmetric keys are maintained

by both parties revokes the need for frequent key refresh. From this point of view, someone could claim that the real communication overhead for simple key updates is practically negligible. The complete sequence of messages for each event is presented in protocol 13.

Protocol 13 The Second MBRA protocol proposed in [48]

Initial Keying:
(1) BS $\Rightarrow$ SS: $h_0, h_1, ..., h_n$
(2) MS $\rightarrow$ BS: $\sigma_{i,j}, R_i, A_i$
(3) BS $\Rightarrow$ SS: $\sum_{j=1}^{n} \sigma_{j,2}, \sum_{j=1}^{n} \sigma_{j,3}, ..., \sum_{j=1}^{n} \sigma_{j,n}$
Key Update:
(1) BS $\Rightarrow$ SS: $h_0, h_1, ..., h_n$
(2) MS $\rightarrow$ BS: $\sigma_{i,j}, R_i, A$
(3) BS $\Rightarrow$ SS: $\sum_{j=1}^{n} \sigma_{j,2}, \sum_{j=1}^{n} \sigma_{j,3}, ..., \sum_{j=1}^{n} \sigma_{j,n}$
Rekeying at Join Event:
(1) BS $\rightarrow$ new SS: $h_2, h_3, ..., h_{n+1}$
(2) BS $\Rightarrow$ SS: h_{n+1}
(3) MS $\rightarrow$ BS: $\sigma_{n+1,j}, R_{n+1}, A_{n+1}$
(4) BS $\Rightarrow$ SS: $\sum_{j=1}^{n} \sigma_{j,2}, \sum_{j=1}^{n} \sigma_{j,3}, ..., \sum_{j=1}^{n} \sigma_{j,n+1}$
Rekeying at Leave Event:
(1) BS $\Rightarrow$ SS: $\sigma_{l,1}, \sigma_{l,1}, ..., \sigma_{l,n}$

From what is described above we can conclude that there are only two solutions ([49], [48]) that manage to cover all security inefficiencies of the existing MBRA algorithm. The solution described in [49] scales poorly especially during initial network setup as it requires many unicasts. On the other hand, the scheme given in [48] is based on asymmetric encryption and may be considered too heavy for some of the today's wireless network implementations. Note that these solutions do inflict modifications on (or totally ignore) the standard MBRA protocol and hence are unable to work along with existing implementations. Table II summarizes the characteristics of the various protocols discussed in this paragraph.

D. Enhancing Security for Mesh Mode

Zhou and Fang suggest some modifications to enhance security in Mesh mode [65]. More specifically, they propose the use of certificates to achieve authentication in link establishment of neighbouring nodes. They introduce the concept of Mesh certificate -an additional certificate- issued by the Authorization center during authorization. Therefore, they recommend a different protocol for link establishment that does not rely on the OSS.

Kwon et al. [51] suggest that a minor modification on the MSH-NCFG: NetEntryOpen message may prevent attackers from masquerading as authorization nodes. This modification involves the addition of the following information $MAC_{BS}|Serial_{BS}|H(MAC_{BS}|Serial_{BS}|AK_{new_node})$ as part of the standard MESH-PKM-RSP message. The authors also propose a modified version of the mesh sub-header which can solve the problem of encryption/decryption of the data forwarded to each network hop. A formal representation of the message flow of this scheme is given in protocol 14.

Protocol 14 Modified Authorization Protocol [51]

1: $Node_A$ mesh certificate, $nonce_A$ encrypted, frame number, ID of $Node_A$, ID of $Node_B$, $Signature_A$
2: $Node_B$ mesh certificate, $nonce_B$ encrypted, frame number, ID of $Node_B$, ID of $Node_A$, $Signature_B$
3: Result, $Signature_A$

After successful run of the protocol above both sides will end up knowing two random numbers ($nonce_A$ and $nonce_B$) which will be used for the production of the link key: $LK_{AB} = H(IDofNode_A, IDofNode_B, nonce_A, nonce_B)$ where $H()$ is a hash function.

In [50] the same authors propose a reputation scheme that defeats these vulnerabilities (of the WiMAX mesh mode) relative to misbehaving nodes. In this scheme the higher the reputation of a given node the easier this node can be selected for sponsor. The reputation value is shared by broadcasts of the MSH-NCFG control message. Also, this protocol employs end-to-end encryption as opposed to link encryption. This is done to assure that intermediate nodes will never be able to have knowledge of the actual data they are assigned to forward.

E. Other Solutions

The work in [29] suggests that certain actions could also be applied in the application layer in order to increase the overall security of a WiMAX network. Such methods include intrusion detection systems, firewalls for gateways, access control to specific applications and the use of session border controllers.

The authors in [66], [67] recommend an interesting way to deal with vulnerabilities of the mobile WiMAX that may lead to DDoS attacks. They propose to use some discarded information such as the upper 64 bits of the HMAC/CMAC (which are normally truncated) as a proof that an MS has already been registered to the network. They name this portion of data as Shared Authentication Information (SAI) and they apply SAI checking, before the normal HMAC/CMAC verification process as a way to relief the system from heavyweight procedures. For example, in a roaming MS scenario, when the MS has to change its serving BS, naturally the Access Services Network Gateway (ASN GW) will have to generate the AK and send it to the new BS. The latter will have to produce the HMAC/CMAC to verify the integrity of the RNG-REC message. On the other hand, having an ASN GW send 64 bits of data to the new BS and then conducting a simple check is a much more lightweight process. The authors illustrate the efficiency of their approach against the secure LU DDoS attack (see section III-B4).

V. Analysis and Discussion

In this section we attempt to validate our allegations regarding the qualitative characteristics of some the attacks described above by providing quantitative assessment. The attacks included in this section were chosen because: (a) it is possible to evaluate them against some quantitative characteristics, (b) their impact is highly bounded to these characteristics, and (c) they are representative for a broader category of attacks. In that way, MOB_TRF-IND Water Torture Attack is indicative for all watertorture attacks, the RNG-REQ DDoS one is quite similar in results with other DDoS attacks, while the AUTH-REQ Replay DoS Attack can be expanded to attacks that are based on imposing computational burden to BS/Network.

TABLE II
COMPARISON OF VARIOUS SCHEMES USED FOR MBS KEY REFRESH

Protocol	Rekey at Expiry	Rekey at Join	Rekey at Leave	Number of Keys	Admin. Complexity	F/B Secrecy	Insider Attacks
MBRA	N Unicasts, 1 Broadcasts	Not Considered	Not Considered	N+1	Low	No	No
[17]	N Unicasts	Not Considered	Not Considered	1	Low	No	Yes
[49]	N Unicasts, 1 Broadcasts	1 Broadcast	1 Broadcast	1	Low	Yes	Yes
[36]	1 Broadcasts	N Unicasts, M Broadcasts	N Unicasts, M Broadcasts	2K	High	Yes	No
[64]	None	1 Broadcast	1 Broadcast	2	Low	No	No
[33](b)	N Unicasts, 1 Broadcasts	Not Considered	Not Considered	N+1	Low	No	Yes
[48]	N+1 Unicasts, 2 Broadcasts	2 Broadcasts	1 Broadcast	2N	Low	Yes	Yes

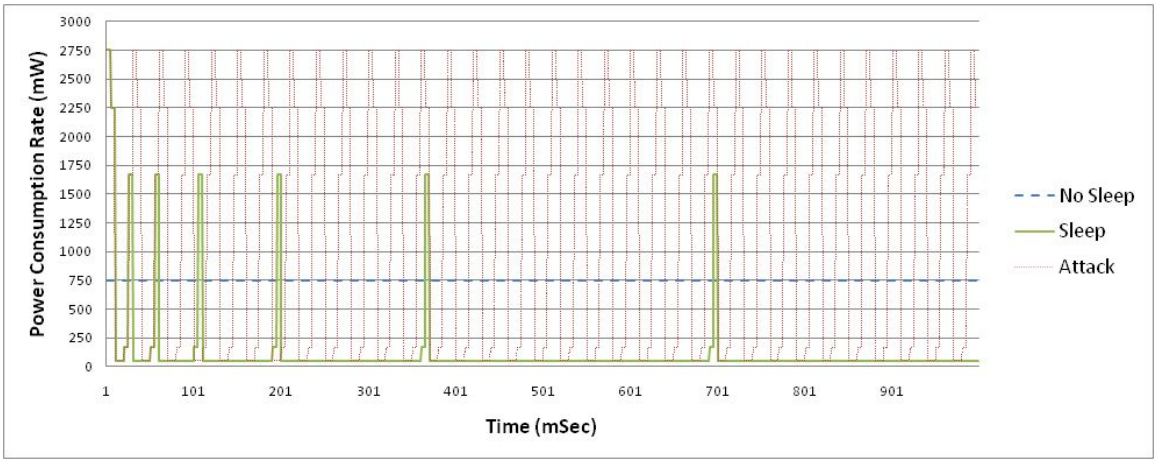


Fig. 9. Snapshot of the energy consumption rate during the first second of operation

A. Analysis of MOB_TRF-IND Water Torture Attack

In order to evaluate the amount of energy consumption caused by attack described in section III-B2, we have proceeded to an analysis which involves the 3 following scenarios: (a) An MS does not support Sleep mode but does not send or receive any traffic for a given period of time, (b) An MS does support Sleep mode and does not receive any traffic for the same period of time, and (c) An MS which supports Sleep mode is under the attack described in section III-B2. The energy consumption for each of the 3 scenarios respectively can be modelled in equations (2), (3) and (4).

In these equations E_{AVG} is the average amount of energy consumed, T_{T_x} is the time required for transmitting a packet, E_A is the energy consumed in awake mode, E_{T_x} is the energy consumed for transmitting a packet, T_{R_x} is the time required for receiving a packet, E_{R_x} is the energy consumed for receiving a packet, T_S^{min} is the smallest possible time window of unavailability, T_S^{max} is the maximum window of availability, E_S is the energy consumed during unavailability interval, T_L is the time window of availability without performing any operation, E_L is the energy consumed during availability interval without performing any other operation, T_{LR_x} is the time required to receive a message during availability interval, and E_{LR_x} is the energy required for receiving a message during availability interval. All time units are counted in msec, while all energy units are counted in mW. Based on (a) the energy values found in [68], (b) the energy values for transmitting and receiving of a popular commercial device, and (c) the time values retrieved from [69], we have conducted our simulation with the following: $T_{T_x} = 5$, $T_{R_x} = 5$, $T_S^{min} = 10$ (2 frames), $T_S^{max} = 5120$ (1024 frames), $T_L = 5$, $T_{LR_x} = 5$ and $E_A = 750$, $E_{T_x} = 2000$, $E_{R_x} = 1500$, $E_S = 50$, $E_L = 170$. In our experiments we assume that the network is operating in OFDMA/TDD with 10Mhz bandwidth. The frame duration is $5msec$ and for simplifying the calculations all three scenarios assume that packets are transmitted to the MS immediately and there is no delay. Also, the energy consumed for other operations of the MS (those relevant to the operating system for example) are neglected. Figure 9 presents a snapshot of the instantaneous current consumed for each of the three scenarios during the first second of operation, while figure 10 illustrates the average energy consumption.

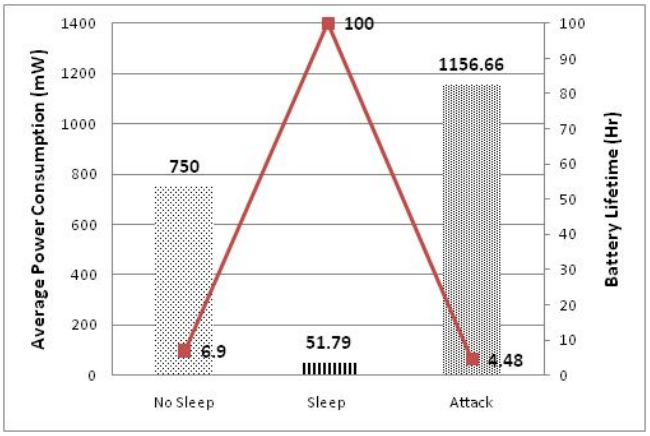


Fig. 10. Average energy consumption Under MOB-TRF-IND-Water-Torture-Attack Attack

The results of the analysis indicate that by unleashing a MOB_TRF-IND Water Torture Attack, the attacker will be able to achieve an energy consumption rate which surpasses that of an MS with no Sleep mode support for over 54%. More specifically, the average power consumption is 750 mW for the

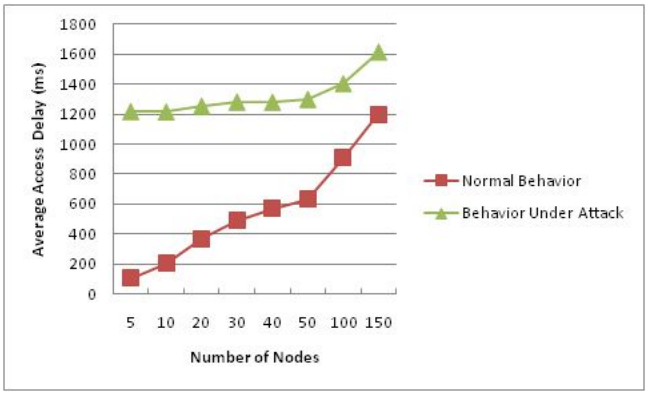


Fig. 11. Average access delay in msec for a different number of contending nodes

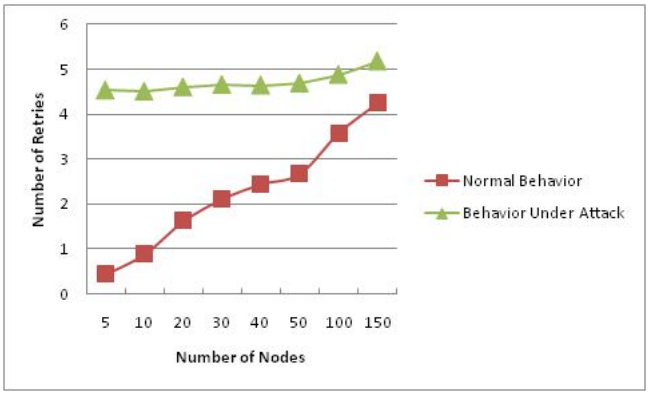


Fig. 12. Average contention retries for a different number of contending nodes

case where the MS does not support Sleep mode (scenario a), only 51.7 mW for the case where Sleep mode is enabled (scenario b), and 1156 mW for this last case, i.e. the attack takes place. We can easily deduce that the energy consumption during the first scenario would deplete a battery of 1400 mAH / 3.7 V (such as the ones equipped by contemporary smartphones) in 6.9 hours. For the second scenario the same battery will be drained in 100 hours while under attack the battery would be depleted in about 4.48 hours. This can prove quite annoying for users of handheld devices for example, while it is doubtful if it will cause disturbance to users of larger energy capacity devices such as laptops. While these values may not reflect realistic discharge rates they are indicative of the impact of the attack.

B. Analysis of RNG-REQ DDoS Attack

For evaluating the impact of RNG-REQ DDoS Attack, the following scenario is considered: A number of MSs which has arrived since the last UCD transmission receives a new UCD message at instance 0 so all MSs are cleared to enter contention for initial ranging process. We consider this simulation for just a time frame as big as the UCD interval (5 sec) but the aggressor unleashes her attack only during the first second of the ranging process. This actually is an interval most likely be chosen in real attack conditions since the Back Off (BO) window size is still small and the collision probability is quite large. During this interval the attacker is transmitting an RNG-REQ message on every single transmission opportunity of every frame. For this simulation scenario we evaluated the initial ranging process in normal operation as well as under attack. More specifically, the behavior of the network in matters of access delay and number of retries is considered under different number of contenting mobile nodes.

For the simulation experiment the following assumptions have been made: frame duration of 5 msec, initial BO window 8, final (maximum) BO window 1024, UCD interval 5 sec, T3 200 msec, simulation duration 5 sec.

The attack causes all contending MSs to collide and as a result to progressively set their backoff window to a very high interval. This has an immediate effect in the access delay thereafter. Still, the total number of RNG-REQ messages transmitted by the attacker in the 1 sec period of attack is not more than 600 messages with total traffic about 96 Kbps (assuming that the RNG-REQ message is 20 bytes). If there is a number of collaborating attackers this value per user can become even smaller. This makes it even harder for deployed defence mechanisms in the BS (such as Intrusion Detection Systems) to become alerted of this abnormality.

Figures 11 and 12 illustrate the delay and number of retries an MS has to make in both scenarios. One can notice that when the attack is unleashed against 5 contending nodes (this corresponds to an arrival rate of 1 node per second) becomes comparable to that of 150 contending nodes (arrival rate of 30 nodes per second) in normal conditions which justifies our classification of this attack as major.

At this point the reader should notice that the TBEB algorithm is also part of the bandwidth request mechanism. Therefore, attacks such as RNG-RSP DoS, Signalling DoS (in the unlikely event of success), PKM-RSP: Auth Invalid DoS, Secure LU DDoS as well as MCA-REQ DoS cause very similar results as the one investigated in this section. In most cases such attacks will force many MSs to disconnect simultaneously. Naturally, after that this large number of MSs will attempt to reconnect performing Initial Network Entry. Eventually, this will result in a large number of MSs contending for a small number of TO in the Initial Ranging step, which is actually the bottleneck of the whole Initial Network Entry process. A very recent paper [70] conducts similar experiments on the ns-2 simulator and concludes that parameters of the Initial Ranging step should be considered critical for system security as a possible inaccurate setting may lead to serious DoS attacks or poor system performance.

C. Analysis of AUTH-REQ Replay DoS Attack

We considered the situation where different number of nodes perform the attack described in section III-E3 against a specific BS. We monitored the amount of CPU load imposed to the system as well as the total amount of time that is required from the BS to serve all the requests. Our purpose is to evaluate the computation burden of this attack and attempt to estimate the amount of client requests needed to (over)stress the BS. The number of Auth-Req considered starts at 100 and scales up to 1000 messages. The maximum amount of Auth-Req messages (1000) reflects the value of 10% of the

$$E_{AVG} = \begin{cases} \frac{T_{T_x}(E_{T_x}+E_A)+T_{R_x}(E_{R_x}+E_A)+T_S^i E_S+T_L E_L+T_{LR_x}E_{R_x}}{T_{T_x}+T_{R_x}+T_S^{min}+T_L+T_{LR_x}}, \text{if } 0 <= i <= n \\ \frac{T_{T_x}(E_{T_x}+E_A)+T_{R_x}(E_{R_x}+E_A)+(T_S^{max}E_S+T_L E_L)+T_{LR_x}E_{R_x}}{T_{T_x}+T_{R_x}+T_S^{min}+T_L+T_{LR_x}}, \text{if } i > n \end{cases} \quad (2)$$

$$E_{AVG} = \frac{T_{T_x}(E_{T_x}+E_A)+T_{R_x}(E_{R_x}+E_A)+T_S^{min}E_S+T_L E_L+T_{LR_x}E_{R_x}}{T_{T_x}+T_{R_x}+T_S^{min}+T_L+T_{LR_x}} \quad (3)$$

$$E_{AVG} = E_A \quad (4)$$

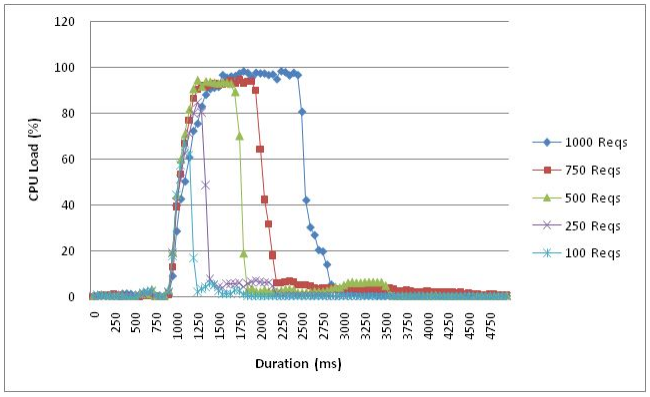


Fig. 13. System CPU load during an Auth-Req Attack

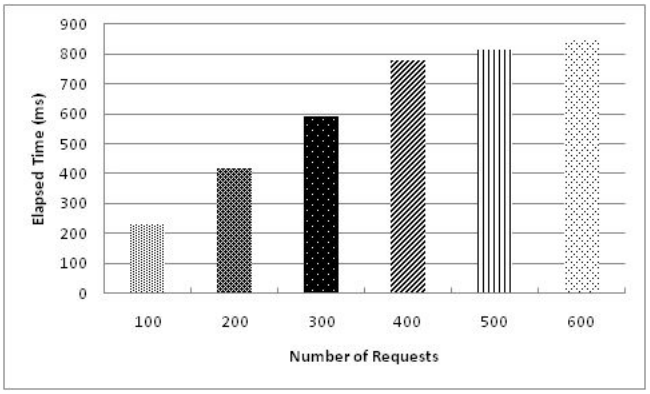


Fig. 14. Total amount of time required to serve all the Auth-Req messages

number of maximum simultaneous connections supported by the state-of-the-art BS equipments today [71]. This, apparently small, percentage is a rather realistic attack condition as the attacker must first eavesdrop and create a database of valid Auth-Req and then make sure that the corresponding CIDs are still active. The experiments were conducted in a custom made simulation environment written in C++ and tested on a Windows 7 (64 bits) Intel Core i7 2.80 Ghz machine incorporating 4 GB of RAM memory. Modern BS equipment is expected to have similar computational power and have analogous performance. Figures 13 and 14 present the CPU load and delay (in terms of service times) respectively.

From the experimental results it is obvious that a significant penalization to the system for a considerable amount of time happens only for more than 500 simultaneous requests. In this case, the CPU load peaked at 94.24% and remained at high levels of 69.3% average for 814 msec. In the case of 1000 simultaneous requests the CPU load peaked at 98.4% and remained at high levels of 77% average for about 2 seconds (2050 msec). Typically, BSs that support a large amount of simultaneous connections are expected to incorporate a stronger CPU than those that support a smaller number of connections. Generalising this empirical study we could conclude that the AUTH-REQ Replay DoS Attack can be fruitful for the attacker only if she is willing to invest time and effort to eavesdrop over a number of Auth-Req messages of at least 5% of the simultaneous connections the victim BS can support.

VI. Conclusions

So far, several research papers have revealed vulnerabilities of the IEEE 802.16 specification. In this paper we attempted to gather and organize documented attacks against this family of standards to facilitate better understanding of their contributions and discuss their merits and shortcomings. Additionally, we evaluated the qualitative characteristic of each attack with reference to the specification in terms of both breadth and depth. After the analysis it becomes clear that many attacks are not possible against the intended version of the standard while most of them can only cause minor damage or commotion to the network. More specifically (excluding the deprecated mesh mode attacks), about 22% of the attacks was evaluated as major, about 11% moderate while -unlike the statements included in the original papers- as much as 52% of the attacks was classified as simply minor. Note that this survey has argued that about 15% of the attacks found in literature are infeasible against the most widely adopted amendment of the standard (802.16e-2005) and its subsequent versions. Table I gathers the attacks described in literature and presents them in organized manner along with their classification.

Excluding the obsolete mesh mode, another interesting conclusion is that about 70% of the attacks aim at causing annoyance to a number of users or DoS to the network. This should be considered by far the most common attack type expected to be found in WiMAX networks. On the other hand, the 3 attacks that might lead to loss of user confidentiality are extremely hard to implement and may prove impractical under realistic conditions (therefore are classified as minor). These percentages along with the categorization of the attacks might prove a useful tool in the field of Intrusion Detection. Since WiMAX networks posses different characteristics than their wired counterparts the compilation of a new dataset which will include the signatures of these attacks will prove of great importance to the research community.

The latest version of the standard, namely 802.16m-2011, has introduced a theoretically more robust security protocol and includes encryption for most of the control messages which in turn is expected to further ameliorate the quality of security provisions. That is, attacks like FPC Water Torture Attack, RES-CMD DoS Attack and DBPC-REQ DoS Attack, to name just a few, will be thwarted. Without doubt, it would be interesting to evaluate the potentiality of the attacks identified by this work under the prism of 802.16m-2011.

APPENDIX

Likehood L : $\mathbf{f_L} = (\mathbf{x}, \mathbf{y}, \mathbf{z})$ **where** $\mathbf{x} \in \mathbf{C}, \mathbf{y} \in \mathbf{D}, \mathbf{z} \in \mathbf{R}$

$$Likely \begin{cases} f_L = (Ex, Ea, Lo) \\ f_L = (Ma, Ea, Lo) \\ f_L = (Ma, Ea, Mo) \\ f_L = (In, Ea, Lo) \\ f_L = (In, So, Lo) \\ f_L = (In, Ea, Mo) \end{cases}$$

$$Possible \begin{cases} f_L = (Ex, So, Lo) \\ f_L = (Ex, Ea, Mo) \\ f_L = (Ex, So, Mo) \\ f_L = (Ex, Ea, Hi) \\ f_L = (Ex, So, Hi) \\ f_L = (Ma, So, Lo) \\ f_L = (Ma, So, Mo) \\ f_L = (Ma, Ea, Hi) \\ f_L = (Ma, So, Hi) \\ f_L = (In, So, Mo) \\ f_L = (In, Ea, Hi) \\ f_L = (In, So, Hi) \end{cases}$$

$$Unlikely \begin{cases} f_L = (Ex, Ha, Lo) \\ f_L = (Ex, Ha, Mo) \\ f_L = (Ex, Ha, Hi) \\ f_L = (Ma, Ha, Lo) \\ f_L = (Ma, Ha, Mo) \\ f_L = (Ma, Ha, Hi) \\ f_L = (In, Ha, Lo) \\ f_L = (In, Ha, Mo) \\ f_L = (In, Ha, Hi) \end{cases}$$

Threat T : $\mathbf{f_T} = \mathbf{f_L} + \mathbf{f_I}$

$$Major \begin{cases} f_I = High \wedge f_L = Likely \\ f_I = High \wedge f_L = Possible \end{cases}$$

$$Moderate \begin{cases} f_I = Medium \wedge f_L = Possible \\ f_I = Medium \wedge f_L = Likely \end{cases}$$

$$Minor \begin{cases} f_I = Low \wedge f_L = Likely \\ f_I = Low \wedge f_L = Possible \\ f_I = Low \wedge f_L = Unlikely \\ f_I = Medium \wedge f_L = Unlikely \\ f_I = High \wedge f_L = Unlikely \end{cases}$$

Impact I : $\mathbf{f_I} = (\mathbf{k}, \mathbf{l}, \mathbf{m})$ **where** $\mathbf{k} \in \mathbf{S}, \mathbf{l} \in \mathbf{T}, \mathbf{m} \in \mathbf{O}$

$$Low \begin{cases} f_I = (Sm, Sh, An) \\ f_I = (Sm, Ln, An) \\ f_I = (Sm, Sh, DoS) \\ f_I = (Me, Sh, An) \\ f_I = (Me, Ln, An) \\ f_I = (Me, Sh, DoS) \end{cases}$$

$$Medium \begin{cases} f_I = (Sm, Ln, DoS) \\ f_I = (Me, Ln, DoS) \\ f_I = (La, Sh, An) \\ f_I = (La, Ln, An) \\ f_I = (La, Sh, DoS) \\ f_I = (Me, Sh, DoS) \end{cases}$$

$$High \begin{cases} f_I = (Sm, Sh, ToS) \\ f_I = (Sm, Ln, ToS) \\ f_I = (Sm, Sh, LoP) \\ f_I = (Sm, Ln, LoP) \\ f_I = (Me, Sh, ToS) \\ f_I = (Me, Ln, ToS) \\ f_I = (Me, Sh, LoP) \\ f_I = (Me, Ln, LoP) \\ f_I = (La, Ln, DoS) \\ f_I = (La, Sh, ToS) \\ f_I = (La, Ln, LoP) \\ f_I = (La, Sh, LoP) \\ f_I = (Ka, Sh, LoP) \\ f_I = (Me, Ln, LoP) \end{cases}$$

REFERENCES

[1] List of Acronyms and Additional Tables (Online Resources) http://www.icsd.aegean.gr/postgraduates/kkolias/attacks-and-countermea -sures-on-802-16/online-resources.pdf

[2] Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Available at: http://standards.ieee.org/getieee802/download/802.11n-2009.pdf

[3] Data-Over-Cable Service Interface Specifications, Baseline Privacy Plus Interface Specification. Available at: http://www.cablelabs.com/specifications/CM-SP-BPI+-C01-081104.pdf

[4] 802.16-2001, I.S. IEEE Standard for Local and Metropolitan

area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems. Availiable at: http://standards.ieee.org/getieee802/download/802.16-2001.pdf.

[5] 802.16-2004, I.S. IEEE Standard for Local and Metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems. Availiable at: http://standards.ieee.org/getieee802/download/802.16-2004.pdf.

[6] 802.16e-2005, I.S. IEEE Standard for Local and Metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems. Availiable at: http://standards.ieee.org/getieee802/download/802.16e-2005.pdf.

[7] 802.16j-2009, I.S. IEEE Standard for Local and Metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems. Available at: http://standards.ieee.org/getieee802/download/802.16j-2009.pdf

[8] 802.16m-2011, I.S. IEEE Standard for Local and Metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems. Available at: http://standards.ieee.org/findstds/standard/802.16m-2011.html

[9] P.W. Chi, Lei, "A Prevention Approach to Scrambling Attacks in WiMAX Networks," World of Wireless, Mobile and Multimedia Networks & Workshops, 2009. WoWMoM 2009. IEEE International Symposium on a , vol., no., pp.1-8, 15-19 June 2009.

[10] Scarfone Karen, Cyrus Tibbs and Matthew Sexton, 2010, "Guide to Securing WiMAX Wireless Communications Recommendations of the National Institute of Standards and Technology", NIST Special Publication 800- 127, Gaithersburg, MD, NIST, Available at: http://csrc.nist.gov/publications/nistpubs/800-127/sp800-127.pdf

[11] M. Barbeau "WiMAX/802.16 Threat Analysis", In Proc. 1st ACM International Workshop on Quality of Service and Security for Wireless and Mobile Networks (Q2SWinet), pages 8-15, October 2005

[12] M. Barbeau, C. Laurendeau, "Analysis of Threats to WiMAX/802.16 Security" in Mobile WiMAX: Toward Broadband Wireless Metropolitan Area Networks, ser. Wireless Networks and Mobile Communications Series, 2007, pp. 347-362.

[13] K. Sansurooah, "An Assessment of Threats of the Physical and MAC Address Layers in WiMAX/802.16"

[14] ETSI. Telecommunications and internet protocol harmonization over networks (TIPHON) release 4; protocol framework denition; methods and protocol for security; part 1: Threat analysis. Technical Specication ETSI TS 102 165-1 V4.1.1, 2003.

[15] T. Han, N. Zhang, K. Liu, B. Tang, Y. Liu, "Analysis of Mobile WiMAX Security: Vulnerabilities and Solutions," Mobile Ad Hoc and Sensor Systems, 2008. MASS 2008. 5th IEEE International Conference on , vol., no., pp.828-833, Sept. 29 2008-Oct. 2 2008.

[16] B. Bhargava, Y. Zhang, N. Idika, L. Lilien, M. Azarmi, "Collaborative Attacks in WiMAX Networks" Security and Communication Networks, 2: 373391.

[17] S. Naseer, M. Younus, A.Ahmed, "Vulnerabilities Exposing IEEE 802.16e Networks to DoS Attacks: A Survey," Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2008. SNPD '08. Ninth ACIS International Conference on , vol., no., pp.344-349, 6-8 Aug. 2008.

[18] J. Hong Kok Han, M. Yusoff Alias, G. Bok Min, "Simulating Denial of Service Attack Using WiMAX Experimental Setup", International Journal of Network and Mobile Technologies (IJNMT), Vol 2, No. 1, 2011, pp. 30-34.

[19] M. Shojaee , N. Movahhedinia , B.T. Ladani, "Traffic Analysis for WiMAX Network Under DDoS Attack," Circuits,Communications and System (PACCS), 2010 Second Pacific-Asia Conference on , vol.1, no., pp.279-283, 1-2 Aug. 2010.

[20] G.Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," IEEE J. Sel. Areas Commun., vol.18, no.3, pp.535-547, Mar 2000.

[21] J, Chi, P. Martins, M. Coupechoux, "A Novel Mechanism for Contention-based Initial Ranging in IEEE 802.16e Networks", 14th Eunice Open European Summer School 2008.

[22] V. V. Girish, V. K. Govindan, S. Baig and V. Yajnanarayana. "A Novel Initial Ranging Algorithm for mobile WiMAX (802.16e)", International Journal of Computer Applications 1(3):95100, February 2010.

[23] L. Lin, W. Jia, B. Han, L. Zhang, "Performance Improvement using Dynamic Contention Window Adjustment for Initial Ranging in IEEE 802.16 P2MP Networks," Wireless Communications and Networking Conference, 2007.WCNC 2007. IEEE , vol., no., pp.1877-1882, 11-15 March 2007.

[24] H. J. Hong Kok, M. Yusoff and M. Goi Bok, "Potential Denial of Service Attacks in IEEE802.16e-2005 Networks", Proc. 9th international conference on Communications and information technologies, pp. 1207-1212, 2009.

[25] F. Ibikunle, "Security Issues in Mobile WiMAX (IEEE 802.16e)", MWS'09 Proc. 2009 IEEE conference on Mobile WiMAX, pp. 117-122, 2009.

[26] J.R. Lee, D.H. Cho, "Performance Evaluation of Energy-Saving Mechanism Based on Probabilistic Sleep Interval Decision Algorithm in IEEE 802.16e," Vehicular Technology, IEEE Transactions on , vol.56, no.4, pp.1773-1780, July 2007.

[27] P.C. Lee, T. Bu, T. Woo, "On the Detection of Signaling DoS Attacks on 3G/WiMAX Wireless Networks", Computer Networks, Volume 53, Issue 15, pp. 2601-2616, 12 October 2009.

[28] B. Kim, J. Park, Y.-H. Choi, "Power Saving Mechanisms of IEEE 802.16e: Sleep Mode vs. Idle Mode," Lecture Notes in Computer Science, vol. 4331, pp. 332-340, 2006.

[29] Rambally Rodney, Abel Vikas Solomon "An Analysis of WiMAX Security Vulnerabilities", International Conference on Wireless Networks and Embedded Systems (WECON 2009) at Chitkara University, Punjab, India 2009.

[30] L. Maccari, M. Paoli, R. Fantacci, "Security Analysis of IEEE 802.16" Communications, 2007. ICC '07. IEEE International Conference on , vol., no., pp.1160-1165, 24-28 June 2007.

[31] Y. Kim H. Lim S., "Shared Authentication Information for Preventing DDoS attacks in Mobile WiMAX Networks," Consumer Communications and Networking Conference, 2008. CCNC 2008. 5th IEEE , vol., no., pp.765-769, 10-12 Jan. 2008.

[32] T Shon, W. Choi, "An Analysis of Mobile WiMAX Security: Vulnerabilities and Solutions" Lecture Notes in Computer Science, 4658, Springer-Verlag, 2007; 8897.

[33] A. Deininger, S. Kiyomoto, J. Kurihara, T. Tanaka, "Security Vulnerabilities and Solutions in Mobile WiMAX", IJCSNS vol.7, no.11, pp.7-15, 2007.

[34] S. Xu, C. Huang, "Attacks on PKM Protocols of IEEE 802.16 and Its Later Versions" Wireless Communication Systems, 2006. ISWCS '06. 3rd International Symposium on , vol., no., pp.185-189, 6-8 Sept. 2006.

[35] A. Altaf, M.Y. Javed, A. Ahmed, "Security Enhancements for Privacy and Key Management Protocol in IEEE 802.16e-2005" Proc. IEEE ACIS Intl Conf. Software Eng., Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD), Aug. 2008.

[36] C. Huang, L. Chang, "Responding to Security Issues in WiMAX Networks", IT Professional 2008; 10(5):15-21.

[37] E. Eren "WiMAX Security Architecture - Analysis and Assessment" Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 2007. IDAACS 2007. 4th IEEE Workshop on , vol., no., pp.673-677, 6-8 Sept. 2007.

[38] T. Shon, B. Koo, J. Park, H. Chang, "Novel Approaches to Enhance Mobile WiMAX Security" EURASIP Journal on Wireless Communications and Networking, vol. 2010.

[39] D. Johnston, J. Walker, "Overview of IEEE 802.16 Security," Security & Privacy, IEEE , vol.2, no.3, pp.40-48, May-June 2004.

[40] A. Mishra, N. Glore, "Privacy and Security in WiMAX Networks", Book Chapter of WiMAX Standarts and Security, CRC Press, 2008.

[41] E. B. Fernandez, M. VanHilst, "An Overview of WiMAX Security" in WiMAX Standards and Security, M. Ilyas, Ed. Boca Raton, FL: CRCPress, 2008, pp. 197204.

[42] S. Vaarala, A. Nuopponen, T. Virtanen, "Attacking Predictable IPsec ESP Initialization Vectors", In Proc. 4th International Conference on Information and Communications Security (ICICS '02), Robert H. Deng, Sihan Qing, Feng Bao, and Jianying Zhou (Eds.). Springer-Verlag, London, UK, UK, 160-172, 2002.

[43] C. B. McCubbin, A. A. Sel, D. P. Sidhu, "Initialization Vector Attacks on the IPsec Protocol Suite. In Proc. 9th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE '00). IEEE Computer Society, Washington, DC, USA, 171-175, 2000.

[44] L. R. Knudsen, J. E. Mathiassen, "A Chosen-Plaintext Linear Attack on DES", In Proc. 7th International Workshop on Fast Software Encryption (FSE '00), Bruce Schneier (Ed.). Springer-Verlag, London, UK, 262-272, 2000.

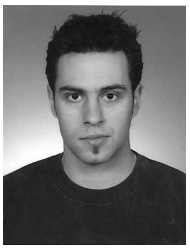
[45] M. Matsui, "Linear Cryptanalysis Method for DES Cipher," Advances in Cryptology, Proceedings Eurocrypt'93, LNCS 765, T. Helleseth, Ed., Springer-Verlag, pp. 386-397 1994,.

[46] L.R. Knudsen, "Block Ciphers Analysis, Design and Applications", Ph.D. Thesis, Arhus University, Denmark, 1994.

[47] Z. Cao, "How to Launch A Birthday Attack Against DES", 2008.

[48] G. Kambourakis, E. Konstantinou, S. Gritzalis, "Revisiting WiMAX MBS security", Computers & Mathematics with Applications, Volume

60, Issue 2, Advances in Cryptography, Security and Applications for Future Computer Science, Pages 217-223, July 2010.
[49] S. Xu, C. Huang, M. Matthews, "Secure Multicast in WiMAX" Journal of Networks, North America, 3 Feb. 2008.
[50] B. Kwon, R. A. Beyah, J. A. Copeland, "Key Challenges in Securing WiMAX Mesh Networks". Security and Communication Networks, 2: 413426.
[51] B. Kwon , C. Lee, P. Chang Yusun, J. A. Copeland, "A Security Scheme for Centralized Scheduling in IEEE 802.16 Mesh Networks" Military Communications Conference, 2007. MILCOM 2007. IEEE , vol., no., pp.1-5, 29-31 Oct. 2007.
[52] X. Sen, H. Chin-Tser, M.M. Matthews, "Modeling and Analysis of IEEE 802.16 PKM Protocols Using CasperFDR", Wireless Communication Systems. 2008. ISWCS '08. IEEE International Symposium on , vol., no., pp.653-657, 21-24 Oct. 2008.
[53] L. Gong, "A Security Risk of Depending on Synchronized Clocks", ACM SIGOPS Operating Systems Review, v.26 n.1, p.49-53, Jan. 1992.
[54] S. Sreejesh, M.P. Sebastian, "A Revised Secure Authentication Protocol for IEEE 802.16 (e)" Advances in Computer Engineering (ACE), 2010 International Conference on , vol., no., pp.34-38, 20-21 June 2010.
[55] M.S. Rahman, M.M.S. Kowsar, "WiMAX Security Analysis and Enhancement" Computers and Information Technology, 2009. ICCIT '09. 12th International Conference on , vol., no., pp.679-684, 21-23 Dec. 2009.
[56] Gilbert S. Vernam, "Cipher Printing Telegraph Systems For Secret Wire and Radio Telegraphic Communications", Journal of the IEEE, Vol 55, pp109115 (1926).
[57] F. Liu, L. Lu, "A WPKI-Based Security Mechanism for IEEE 802.16e" Wireless Communications, Networking and Mobile Computing, 2006. WiCOM 2006.International Conference on , vol., no., pp.1-4, 22-24 Sept. 2006.
[58] S.S. Hasan, M.A. Qadeer, "Security concerns in WiMAX," Internet, 2009. AH-ICI 2009. First Asian Himalayas International Conference on , vol., no., pp.1-5, 3-5 Nov. 2009.
[59] J. Al-Zaabi, N. Chilamkurti, S. Zeadally, J. Kim, "A Proposed Authentication Protocol for Mobile Users of WiMAX Networks," Human-Centric Computing (HumanCom), 2010 3rd International Conference on , vol., no., pp.1-6, 11-13 Aug. 2010.
[60] J. Huixia, L. Tu, G. Yang, Y. Yang, "An Improved Mutual Authentication Scheme in Multi-Hop WiMAX Network," Computer and Electrical Engineering, 2008. ICCEE 2008. International Conference on , vol., no., pp.296-299, 20-22 Dec. 2008.
[61] H. Tian, L. Pang , Y. Wang, "Key Management Protocol of the IEEE 802.16e", Wuhan University Journal of Natural Sciences.2007, 12(1):59-62.
[62] M. Bellare, P. Rogaway, "Entity Authentication and Key Distribution", Advances in Cryptology CRYPTO 93, pp. 232-249, vol. 773, 1994.
[63] R. Li, Z. Fang, P. Xu, W. Xiao, W. Wang, "Experimental Research on a New Authentication Protocol for Wireless Communication Network Based on WiMAX," Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference on , vol., no., pp.1-4, 12-14 Oct. 2008.
[64] J. Brown, X. Du, M. Guizani, "Efficient Rekeying Algorithms for WiMAX Networks". Security and Communication Networks, 2: 392400.
[65] Y. Zhou, Y. Fang "Security of IEEE 802.16 in Mesh Mode", Military Communications Conference, 2006. MILCOM 2006. IEEE , vol., no., pp.1-6, 23-25 Oct. 2006.
[66] Y. Kim, S. Bahk, "Enhancing Security Using the Discarded Security Information in Mobile WiMAX Networks," Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE , vol., no., pp.1-5, Nov. 30 2008-Dec. 4 2008.
[67] Y. Kim, H. Lim; S. Bahk, "Shared Authentication Information for Preventing DDoS attacks in Mobile WiMAX Networks," Consumer Communications and Networking Conference, 2008. CCNC 2008. 5th IEEE , vol., no., pp.765-769, 10-12 Jan. 2008.
[68] Kwanghun Han, Sunghyun Choi "Performance Analysis of Sleep Mode Operation in IEEE 802.16e Mobile Broadband Wireless Access Systems," Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd , vol.3, no., pp.1141-1145, 7-10 May 2006.
[69] 802.16-2009 - IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems Available at: http://standards.ieee.org/getieee802/download/802.16-2009.pdf
[70] Juan Deng, Richard R. Brooks, and James Martin, "Assessing the Effect of WiMAX System Parameter Settings on MAC-level Local DoS Vulnerability", International Journal of Performability Engineering, 8 (2) 163-178, (2012).
[71] Huawei Technologies Co., Ltd. Huawei WiMAX Base Station3703. http://www.huawei.com/le/download.do?f=2682



Constantinos Kolias was born in Athens, Greece in 1982. He holds a Diploma in Computer Science from Technological Educational Institute of Athens, Greece and MSc in Information and Communication System Security. He is currently a Ph.D. candidate, supervised by Dr. G. Kambourakis, at the Department of Information and Communication Systems Engineering, University of the Aegean, Greece. His primary research interests lie in the field of Mobile Ad-Hoc Networks (MANet) Security, Wireless Sensor Network (WSN) Security, Peer-to-Peer (P2P) Networks Security, Ubiquitous Computing, Pervasive Applications Development, User Adaptive Applications Development.

Georgios Kambourakis was born in Samos, Greece, in 1970. He received the Diploma in Applied Informatics from the Athens University of Economics and Business (AUEB) in 1993 and the Ph.D. in information and communication systems engineering from the department of Information and Communications Systems Engineering of the University of Aegean (UoA). He also holds a M.Ed. from the Hellenic Open University. Currently Dr. Kambourakis is an Assistant Professor at the Department of Information and Communication Systems Engineering of the University of the Aegean, Greece. His research interests are in the fields of Mobile and ad-hoc networks security, VoIP security, security protocols, Public Key Infrastructure and mLearning and he has more than 80 publications in the above areas. He has been involved in several national and EU funded R&D projects in the areas of Information and Communication Systems Security. He is a reviewer of several IEEE and other international journals and has served as a technical program committee member in numerous conferences. Dr. Kambourakis is a member of the Greek Computer Society.

Stefanos Gritzalis is a Professor at the Department of Information and Communication Systems Engineering, University of the Aegean, Greece and the Director of the Laboratory of Information and Communication Systems Security (Info-Sec-Lab). He also serves as the Special Secretary at the Greek Ministry of Public Administrative Reform and Electronic Governance. He holds a BSc in Physics, an MSc in Electronic Automation, and a PhD in Information and Communications Security from the Dept. of Informatics and Telecommunications, University of Athens, Greece. He has been involved in several national and EU funded R&D projects. His published scientific work includes 30 books or book chapters, 90 journals and more than 120 international refereed conference and workshop papers. The focus of these publications is on Information and Communications Security and Privacy. His most highly cited papers have more than 1.200 citations. He has acted as Guest Editor in 22 journal special issues, and has been involved in more than 30 international conferences and workshops as General Chair or Program Committee Chair. He has served on more than 250 Program Committees of international conferences and workshops. He is an Editor-in-Chief or Editor or Editorial Board member for 15 journals and a Reviewer for more than 45 journals. He has supervised 10 PhD dissertations. He was an elected Member of the Board (Secretary General, Treasurer) of the Greek Computer Society. His professional experience includes senior consulting and researcher positions in a number of private and public institutions. He is a Member of the ACM, the IEEE, and the IEEE Communications Society "Communications and Information Security Technical Committee.