Security Advances and Challenges in 4G Wireless Networks

N. Seddigh, B. Nandy, R. Makkar
Solana Networks
Ottawa, Canada

J.F. Beaumont
Defence Research & Development Canada
Ottawa, Canada

Abstract — **This paper presents a study of security advances and challenges associated with emergent 4G wireless technologies. The paper makes a number of contributions to the field. First, it studies the security standards evolution across different generations of wireless standards. Second, the security-related standards, architecture and design for the LTE and WiMAX technologies are analyzed. Third, security issues and vulnerabilities present in the above 4G standards are discussed. Finally, we point to potential areas for future vulnerabilities and evaluate areas in 4G security which warrant attention and future work by the research and advanced technology industry.**

Index Terms— **Security, 4G Wireless, LTE, WiMAX**

I. Introduction

The next generation of wireless technology (fourth generation or 4G) is intended to support broadband performance and enable voice/video multimedia applications. The enabling technologies and standards for 4G wireless communications allow for significant increases in data rates over 2G (second generation), 3G (third generation) and 3.5G wireless technologies. At the present time, LTE (Long Term Evolution) and WiMAX (Worldwide Interoperability for Microwave Access) are the two technologies considered as candidates to achieve the 4G wireless performance objectives.

4G wireless technologies have a number of key differences in comparison with 3G and earlier wireless technologies. A key defining factor is the fact that 4G wireless networks will operate entirely based on the TCP/IP architecture and suite of protocols. This design decision is intended to drive costs down as IP has become the ubiquitous choice for data networking across heterogeneous technologies. It will also result in opening up what was previously a closed cellular market restricted to relatively few vendors.

However, one consequence in moving to an open set of communication protocols (TCP/IP suite) is the anticipated increase in security issues compared with 2G and 3G. Significant attention has been given to security design during the development of both the LTE and WiMAX standards. However, due to the open nature and IP-based infrastructure for 4G wireless, it appears that further significant attention needs to be given to understand and study the security issues.

The task of securing 4G wireless networks and systems is a challenging one. The designers of 4G wireless systems have a plethora of security technologies and previous wireless security experiences to draw upon. Key issues faced by designers of 4G wireless security include the following: (i) Security issues for 4G mobile wireless devices and the supporting network architectures will need to take into account all the security issues of accessing the Internet from a fixed location plus the additional requirements for flexibility and mobility (ii) Every time additional cryptographic methods and security mechanisms are applied to IP networks, there is an impact on the performance and traffic handling capacity of the service provider's network. Standards bodies and vendors will need to weigh the security risk against the performance/costs of a particular security solution (iii) A new generation of 4G devices and applications are sure to emerge within the next decade. All these applications and devices will need to be protected from a growing range of security threats.

This paper studies security advances and challenges in 4G technologies LTE and WiMAX through a review of state of the art research. The intent is to identify areas in 4G security which warrant attention and future work by the research community and advanced technology industry. We believe this is one of the first publications focusing on both LTE and WiMAX security. We focus primarily on specific MAC layer security issues that are unique to LTE and WiMAX.

The remainder of this paper is organized as follows. Section II and III discuss 4G technology standards and evaluate current 4G network architectures. Section IV analyzes the evolution of security in wireless standards. Section V reviews the 4G security architecture while Section VI focuses on security issues in 4G wireless.

II. 4G Wireless Standards

Previously, the ITU (International Telecommunications Union) defined the IMT-2000 (International Mobile Telecommunications-2000) standard as a global standard for 3G wireless communications. More recently, the ITU embarked on an initiative to define a wireless system beyond IMT-2000 – referred to IMT-Advanced. Work is still underway on standards definition for IMT-Advanced – which may eventually be considered as the formal specification for 4G wireless. The ITU framework and overall objectives for wireless systems beyond IMT-2000 considers both the radio access network (RAN) and the "core network" [1].

4G wireless technology must support the following criteria: (a) high data rate (1Gps peak rate for low mobility and 100Mbps peak rate for high mobility) (b) high capacity (c) low cost per bit (d) low latency (e) good quality of service (QoS) (f) good coverage and (g) mobility support at high speeds. In 4G, much effort has been invested in bandwidth optimization and efficiency gain techniques. Flexible

bandwidth allocation schemes and modulation approaches at the base stations allow for differing service levels depending on the capability of the end user device. 4G wireless will also be distinguished by the fact that voice and data will be carried on the same infrastructure utilizing the same set of network protocols.

Additionally, the technology should provide a clear evolutionary path to the ITU IMT-Advanced standard for 4G mobile wireless. Several new broadband wireless access technologies have been developed by standards bodies such as 3GPP, 3GPP2, IEEE802.16 & the WiMAX Forum to offer mobile broadband wireless access. Some of these technologies have been labeled as sufficient to meet the requirements for 4G wireless. Initially, candidate technologies for 4G wireless standard included: (i) HSPA+ (High Speed Packet Access) (ii) UMB (Ultra Mobile Broadband) (iii) LTE (iv) Mobile WiMAX (v) XGP(eXtended Global Platform)

In reality, only 3 of the 5 technologies were seriously considered as candidates for the 4G wireless standard. HSPA+ provides an upgrade over 3G wireless and allows download speeds of 14Mbps to 42Mbps. Many wireless carriers are in the process of evolving their networks to HSPA+ because it requires less investment and time than upgrading to 4G technologies such as LTE. However, HSPA+ technology will not meet the ITU IMT-Advanced requirements of an all-IP interface. Despite the significant bandwidth and performance improvements, it did not meet the stringent ITU requirements for 4G. As a result, the 3GPP-Release 8 LTE standard became the evolutionary path towards 4G for UMTS/HSPA (Universal Mobile Telecommunications System).

Another technology under consideration for 4G wireless was the Qualcomm-driven UMB standard. Due to what appeared to be business as opposed to technology related factors, Qualcomm decided to stop working on UMB. XGP is yet another technology described as meeting the 4G wireless requirements. XGP is a 4G upgrade version of the 2G PHS wireless technology which was used by almost 100 million wireless subscribers in Japan/China. While it is possible that XGP may gain momentum within Japan, Asia and a few other countries, PHS represents less than 2% of the global wireless subscriber market. Accordingly, it is not a strong candidate for a global 4G wireless standard.

As a result, LTE and Mobile WiMAX remain as the key wireless 4G technologies.

III. 4G Network Architecture

A. WiMAX

The Figure below illustrates the end-to-end network architecture for mobile WiMAX. It consists of two key entities: (i) Access Services Network (ASN) and (ii) Connectivity Services Network (CSN). The core elements in the ASN are the base station (BS) and ASN gateway (ASN-GW) which are connected over an IP infrastructure. The ASN-GW provides security anchoring, traffic accounting and mobility support for the mobile station (MS). The mobile IP home agent (HA) in the CSN enables global mobility.

There are a number of key elements in the operation of the WiMAX network architecture. First, the AAA (Authentication, Authorization, and Accounting) server located in the CSN network processes control signals from the ASN-GW to authenticate the MS against the MS's profile stored in the AAA server's database. Once authenticated, the AAA server sends the MS's profile including the QoS parameters to the ASN-GW. The Home Agent (HA) processes control signals from the ASN-GW and assigns a Mobile IP address to the MS and anchors the IP payload. The HA server provides connectivity to the Internet for data traffic.

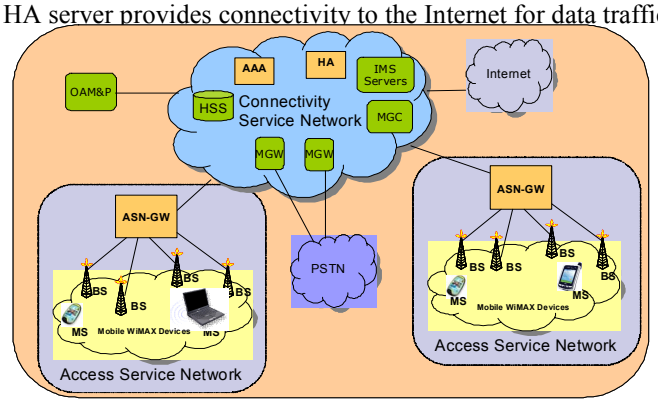

Figure 1: Mobile WiMAX Network

If the MS makes a VoIP call, control is passed to the CSN IMS (IP Multimedia System) servers which then process the call. If the call is to a telephone number that is outside the WiMAX network, the IMS servers select the appropriate Media Gateway Controller (MGC) / Media Gateway (MGW) to interface to the Public Switched Telephone Network (PSTN). Alternatively, if the call is to an end unit in another 3GPP or 3GPP2 network, it is routed through the Inter-working Gateway Unit within the CSN.

Mobile Stations (MS) communicate with Base Stations (BS) using the 802.16e (802.16m in the future) air interface. The communication between the MS and BS is via an all-IP bearer and control. WiMAX does not have a TDM (Time Division Multiplexing) bearer. Like LTE, it is an all-IP flat network. WiMAX MS user traffic is tunneled as payload between the BS and ASN-GW.

In most service provider configurations, the CSN network elements are redundant and geographically separate. The ASN-GW network elements within the ASN are also configured in a redundant manner typically within the same premises. A Network Access Provider (NAP) can have multiple ASNs. Mobility within these ASNs does not have to be anchored at the CSN. Roaming is supported when the MS roams out of its Home Network Service Provider (NSP) to a Visited NSP. In such cases, the AAA server in the Visited NSP uses control signalling to obtain the credentials and profiles from the Home NSP. Bearer traffic is not sent from the visited NSP to the home NSP. Various mobility scenarios are supported including intra-ASN-GW, inter-ASN-GW and anchored CSN mobility. When the MS moves from one BS to another BS served by the same ASN-GW, calls can be switched seamlessly using signaling.

B. LTE

Figure 2 below depicts the LTE high level architecture [2]. The UE (user equipment) such as smart phones or laptops connect to the wireless network through the eNodeB within the E-UTRAN (Evolved UMTS Terrestrial Radio Access Network). The E-UTRAN connects to the EPC (Evolved Packet Core) which is IP-based. The EPC connects to the provider wireline IP network.

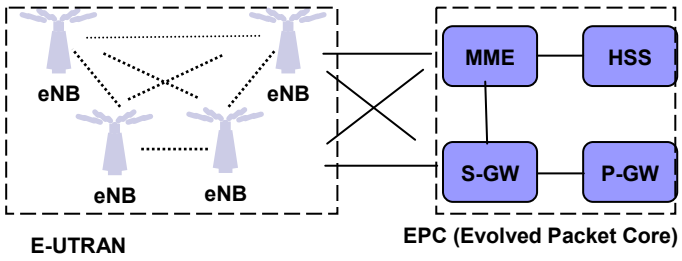

Figure 2: LTE – System Architecture Evolution (SAE)

In comparison with 3G wireless, the LTE network architecture has a number of key differences. First, it has fewer types of network elements (NEs). An LTE network consists of 2 types of NEs: (i) the eNodeB which is an enhanced base station (ii) the Access Gateway (AGW) which incorporates all the functions required for the EPC. Second, LTE supports a meshed architecture which allows greater efficiency and performance gains. For example, a single eNodeB can communicate with multiple AGWs. Third, a flat all IP-based architecture is utilized. Traffic originating at a UE is generated in native IP format. These packets are then processed by the eNodeB and AGW using many of the standard functions that are present in IP-based devices such as routers. In addition, the signalling and control protocols for the network are IP-based.

The eNodeB (eNB) is the single type of system in the 4G E-UTRAN - it incorporates all the radio interface-related functions for LTE. The AGW is the single type of system in the LTE EPC. The eNB communicates with the UE as well as with the AGW in the EPC. The communication with the AGW occurs over the transport network. Some of the other high level functions carried out by the eNB include (i) inter-cell radio resource management (RRM) (ii) Radio admission control (iii) Scheduling via dynamic resource allocation (iv) Enforcement of negotiated QoS on uplink (v) compression/decompression of packets destined to/from the UE. The AGW consists of multiple modules including the (i) HSS (Home Subscriber Server) (ii) the P-GW (Packet Data Network Gateway) (iii) the S-GW (Serving Gateway) and (iv) the MME (Mobility Management Entity). The LTE standard has sufficient flexibility to allow vendors to combine these different modules into a single device or into multiple devices. e.g separating the MME and S-GW into different devices.

UE data packets are backhauled from the eNB to the AGW over the provider's transport network using IP and MPLS (Multiprotocol Label Switching) networks as the primary vehicle for backhaul in 4G.

The MME is the key control-node for the LTE. It is responsible for managing the UE identity as well as handling mobility and security authentication. It tracks the UE while it is in idle mode. The MME is responsible for choosing an S-GW for a UE during its initial attach to the network as well as during intra-LTE handover. The MME authenticates the user via interaction with the HSS. The MME also enforces UE roaming restrictions. Finally, the MME handles the security key management function in LTE.

The S-GW terminates the interface towards the E-UTRAN. It has key responsibility for routing and forwarding data packets. It acts as the mobility anchor during inter-eNB handovers. The S-GW also has a mandate to replicate packets to satisfy lawful intercept requirements and functions.

The P-GW terminates the interface towards the packet data network. i.e the service provider wireline network. It is the gateway that ultimately allows the UE to communicate with devices beyond the service provider main IP network. UEs may simultaneously connect to multiple P-GWs in order to connect to multiple provider IP networks. Other key functions carried out by the P-GW include (i) Policy enforcement (ii) Per-user packet filtering (iii) billing & charging support (iv) Anchor for mobility between 3GPP and non-3GPP technologies such as WiMAX and CDMA-based 3G (v) Allocating the IP address for the UE.

The HSS maintains per-user information. It is responsible for subscriber management as well as for security. The HSS contains the subscription-related information to support network entities handling the calls/sessions. The HSS generates authentication data and provides it to the MME. There is then a challenge-response authentication and key agreement procedure between the MME and the UE. The HSS connects to the packet core based on the IP-based Diameter protocol and not the SS7 (Signaling System Number 7) protocol used in traditional telecommunication networks.

IV. Security Evolution In Wireless

A. WiMAX – Security Evolution

The IEEE 802.16 (WiMAX) Working Group wanted to avoid the well known and documented security design issues with IEEE 802.11 by incorporating a pre-existing standard into IEEE 802.16. However, as the standard evolved from 802.16 to 802.16a to 802.16e, the requirements evolved from line of sight to mobile WiMAX. As a result, the security requirements and standards also evolved in order to address changing needs.

The security features introduced in the initial IEEE 802.16 standard have been greatly enhanced in the IEEE 802.16e-2005 standard. Key new features include: (i) PKMv2 (Privacy Key Management version 2) protocol (ii) Message authentication is performed using the HMAC/CMAC (Hash-based Message Authentication Code or Cipher-based Message Authentication Code) scheme (iii) Device/user authentication is carried out using Extensible Authentication Protocol (EAP) methods and (iv) confidentiality is achieved using AES (Advanced Encryption Standard) based encryption.

Despite the above, the security strength of 802.16e required

improvements. The work-in-progress 802.16m standard will further strengthen WiMAX security. Over-the-air security remains a key part of ensuring end-to-end network security in WiMAX. While security architectures have been developed to mitigate against threats over-the-air, there are still a number of associated challenges which are discussed later in this paper.

Our analysis indicates that the industry's main challenge will be to balance security needs with the cost of implementation, performance, and interoperability. In addition, since WiMAX utilizes IP as its transport mechanism for handling control/signalling and management traffic, network operators will have to defend against general IP related security threats as well.

B. LTE – Security Evolution

Security architecture in cellular networks have evolved continually. In 1G (first generation) wireless, intruders could eavesdrop on conversations and gain fraudulent access to the network [3]. In 2G GSM (Global System for Mobile Communications), authentication algorithms were not very strong. A few million interactions with a SIM card could disclose the master security key [4]. In 3G wireless (3GPP-based), the authentication mechanism was enhanced to become a two-way process. Both the mobile device and the network achieved mutual authentication. In addition, 128-bit encryption and integrity keys were utilized to create stronger security [5]. Finally, mechanisms were introduced to ensure freshness of the cipher/integrity keys. As a result, if a security key is compromised or broken, the damage is limited only to the period of time for which the key is valid – instead of having long lasting effects [6].

In 4G LTE, further security improvements were introduced over 3GPP. For example, further layers of abstraction were added in terms of the unique identifiers (ID) for an end-mobile device (UE). In 2G, a solitary unique ID was used on the SIM card, in 3G and subsequently 4G LTE, temporary ID and further abstraction was used so that smaller windows of opportunity exist for intruders to steal identities. Another mechanism to strengthen security in 4G was to add secure signalling between the UE and MME (Mobile Management Entity). Finally, security measures were put in place for inter-working between 3GPP networks and trusted non-3GPP users – using for example, the EAP-AKA (UMTS Authentication and Key Agreement) protocol [7].

Security protection has witnessed significant advances during the evolution from 1G to 4G. However, our analysis indicates that the dual phenomenon of utilizing an open IP-based architecture as well as the sophistication of security hackers means that security issues remain a matter of key concern in 4G systems. Careful attention needs to be given to analyzing security challenges in 4G wireless and rapid development of solutions for threat detection and mitigation.

V. 4G Wireless Security Architecture and Design

From a broad perspective, the security architecture for 4G systems should meet the following security requirements: (i) increased robustness over 3G (ii) user identity confidentiality (iii) strong authentication of user and network (iv) data integrity (v) confidentiality and (vi) inter-working of security across other radio networks.

A. WiMAX – Security Architecture and Design

The IEEE 802.16 standard defines the medium access control (MAC) layer for the wireless link between a BS and a SS/MS. The MAC layer consists of a further sub-layer – the MAC Security sub-layer. This sub-layer handles (i) authentication and authorization (ii) key management/ distribution, and (iii) encryption. Figure 3 captures the WiMAX security protocol stack and depicts many components of the MAC security sub-layer.

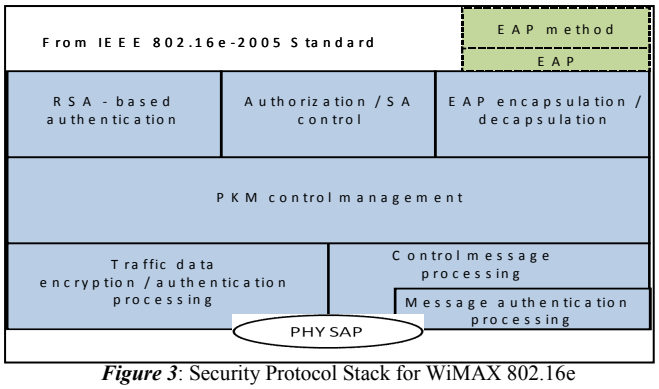

Figure 3: Security Protocol Stack for WiMAX 802.16e

1) Authentication & Authorization

The role of authentication is to verify the identity of a device wishing to connect to the wireless network. WiMAX supports three kinds of authentication: (i) RSA-based authentication (Rivest, Shamir and Adleman public key cryptography) (ii) EAP-based (Extensive Authentication Protocol) authentication (iii) RSA-based authentication followed by EAP-based authentication.

Prior to first use by a user, WiMAX devices require credentials (X.509 digital certificates) to be loaded on the device and also to be programmed in the home network's AAA server. The X.509 certificate used in RSA-based authentication is issued by the SS (subscriber station) manufacturer and contains the SS's public key (PK) and its MAC address. When requesting an authorization key (AK), the SS transmits its certificate to the BS who validates it and then uses the PK to encrypt an AK and transmit it to the SS.

In EAP-authentication, the SS is authenticated by a X.509 certificate or by a unique operator-issued credential such as SIM, USIM or userid and password. The WiMAX standard allows for use of any of three EAP-authentication schemes: EAP-AKA (Authentication and Key Agreement), EAP-TLS (Transport Layer Security) and EAP-TTLS MS-CHAP v2 (Tunneled Transport Layer Security with Microsoft Challenge-Handshake Authentication Protocol version 2). EAP-TTLS is used to support establishment of secure connections in a roaming environment and protection of user credentials.

In general, for authentication and authorization, once the

MS request registration with the BS, the BS communicates with the MS's home AAA server using EAP running over the RADIUS or DIAMETER protocol. On verifying the identity of the MS, the AAA server returns an intermediate key (MSK or Master Session Key) to the authenticator in the MS visiting network. Ultimately, the MSK is translated into another intermediate key – the PMK (Pairwise Master Key). The AK is then generated from the PMK.

2) Key Management

Once the identity of the SS/MS has been validated, traffic keys are then exchanged. IEEE 802.16e defines Privacy Key Management (PKM) for secure key distribution between the MS and the BS. In addition to ensuring synchronization of keying data between the BS and SS/MS, the protocol is used by the BS to authorize SS/MS access to the network.

The IEEE 802.16e standard supports two versions of the PKM protocol: (i) PKM version 1 (PKMv1) which provides a basic set of functionality and (ii) PKM version 2 (PKMv2) which incorporates a number of enhancements. Both versions of the PKM protocol facilitate two functions (i) Access Control - the protocol provides the means for the BS to authenticate an SS/MS. It also provides the BS with the capability to authorize the SS/MS access to the network and any subscribed services. In addition, the protocol enforces periodic re-authentication and reauthorization and (ii) Key Management - the protocol enables the secure exchange of key information between the BS and the SS/MS.

Key management is implemented using a client-server model in PKM. The SS/MS (PKM client) requests keying material from the BS (PKM server). The client receives keying material for services (SAs) that it is authorized to access.

WiMAX communication is secured through the use of five kinds of keys: (i) Authorization key (AK) (ii) Key encryption key (KEK) (iii) Downlink hash function-based message authentication code (HMAC) key (iv) Uplink HMAC key, and (v) Traffic encryption key (TEK).

PKMv2 addresses the requirement for mutual authentication between the SS/MS and BS. It includes new security features such as support for (i) a new key hierarchy for AK derivation and (ii) the Extensible Authentication Protocol (EAP). Mutual authentication allows the BS to validate the identity of the SS/MS and also allows the SS/MS to validate the identity of the BS. PKMv2 supports both an RSA-based authentication process and an EAP-based process.

Similar to PKMv1, the RSA authentication process in PMKv2 involves a three-message exchange. However, the content of the last two messages in the sequence have changed. In addition, instead of the AK being generated by the BS and sent encrypted to the SS/MS, the AK is now generated by the SS/MS using an encrypted pre-PAK (pre Primary Authorization Key). The SS/MS takes the pre-PAK and generates a PAK. The PAK is then used to generate the AK. The BS follows the same process to create the AK. The two generated AKs must match for subsequent communications between the SS/MS and BS to succeed.

The authentication process begins when the SS/MS sends an Authentication Information message containing the X.509 certificate of the SS/MS manufacturer to the BS, as in PKMv1. Immediately following this message, an Authorization Request message is sent by the SS/MS, as in PKMv1. However, the PMKv2 message has an additional random number generated by the SS/MS. Once the SS/MS identity has been validated by the BS, the BS replies with an Authorization Reply message containing the following: (i) the random number generated by the SS/MS and received in the Authorization Request message (ii) a random number generated by the BS (this number is used to verify the freshness of the current message) (iii) the pre-PAK which is RSA encrypted using the SS/MS public key (the pre-PAK will be used by the SS/MS to generate the AK) (iv) the key lifetime of the PAK (v) the PAK sequence number (vi) the SAID list (vii) the X.509 certificate for the BS (viii) the RSA signature of the BS.

With the information contained in the Authorization Reply message, the SS/MS can verify the identity of the BS and generate the AK.

3) Encryption

Only after the successful exchange of keys can encrypted data be transmitted via the WiMAX connection. WiMAX security includes an encapsulation protocol for securing data across the wireless link. The BS uses the AK to generate the TEK. The TEK is utilized for secure encryption of data across the wireless link. Other keys are generated by the BS to facilitate a secure three-way handshake in transmitting the TEK to the SS/MS.

B. LTE – Security Architecture and Design

LTE security requirements cover three levels: (i) Level I which protects communication between the UE and E-UTRAN or MME (ii) Level II which provides protection between elements in the wireline network and (iii) Level III which provides secure access to the mobile station. In general, when compared to 3G, LTE security has (i) Extended authentication and Key Agreement (ii) More complex key hierarchy (iii) More complex inter-working security (iv) Additional security for the eNB (compared to 3G base stations) [8].

1) Key Building Blocks

a) Key Security & Hierarchy

LTE utilizes 5 different keys, each used for a specific purpose and valid only for certain duration. Different keys are used for communication in the E-UTRAN and the EPS. We believe this approach greatly reduces the effect of any possible security compromise. All the keys are derived using the Key Derivation Function (KDF) [9]. The 5 critical security keys derive their basis from the K key with a number of intermediate keys utilized as well. K is the permanent key stored on the USIM (Universal Subscriber Identity Module)

on the UE. CK and IK are the pair of keys derived on the USIM during an AKA exchange. Subsequently, the K_{ASME} key is derived from the CK, IK and SN identity using a KDF [9].

The 5 keys are: (i) $KNAS_{int}$ and $KNAS_{enc}$ integrity and encryption keys respectively are used to protect NAS traffic between the UE and MME (ii) KUP_{enc} key is used to encrypt user data traffic between the UE and eNodeB (iii) $KRRC_{int}$ and $KRRC_{enc}$ integrity and encryption keys respectively are used to protect RRC (Radio Resource Control) traffic between the UE and the eNodeB.

b) Authentication, Encryption & Integrity Protection

Authentication, encryption and integrity protection procedures in LTE focus on the following: (i) Freshness - The authentication vector which is at the heart of the authentication procedure is guaranteed to be fresh. i.e not previously utilized. This is achieved via the sequence numbers exchanged in the messages that serve as input to the ciphering and integrity algorithms [10] (ii) Security algorithms - The algorithms used in the HE (home environment) and USIM to compute the authentication vectors are mostly one-way mathematical functions, where the output is obtained with a given set of inputs, using a pre-defined algorithm. Thus, it is extremely complex for an attacker to try to obtain the inputs using the outputs [10] (iii) Use of IPSEC – The IPSEC protocol is utilized to ensure confidentiality of user traffic as it is transmitted between nodes in the LTE EPS (Evolved Packet System). In addition, IPSEC tunnels are utilized for communication between various nodes in the visited and home networks – for mobile nodes. This introduces a requirement to have S-GW and MME nodes with sufficient processing power to handle encryption and decryption at required speeds so that performance is not hampered [11].

c) Key Management

LTE key management functions include (i) key establishment (ii) key distribution and (iii) key generation. A secure key management mechanism and protocol is required. Otherwise, keys can be leaked, thus rendering established ciphering and integrity mechanisms irrelevant. At the same time, especially in a wireless all-IP network, it is important for the mobile network to operate with fast handover and security so that there is no impact on perceived user quality [12].

LTE, like its predecessors UMTS and GSM, utilize the AKA (Authentication and Key Agreement) procedure for key establishment and verification. In particular, it uses the EPS-AKA procedure. AKA has three stages: (i) Initiation (ii) Transfer of credentials (iii) Challenge-response exchange. During the initiation stage, the mobile UE provides the network with its identity – either IMSI (International Mobile Subscriber Identity) or TMSI (Temporary mobile subscriber identity). Based on this identity, the network initiates the authentication procedure [3].

d) Unique User Identifiers

A key aspect of LTE security is to protect malicious intruders from learning about or misappropriating the identity of mobile users. In particular, the unique identifier for the mobile user is a number that if compromised can result in a number of security threats including (a) tracking and profiling the user's movements (b) illegal access to the network (c) denial of service attacks

In LTE, a number of counter measures have been developed to address the above issues [10]. The major strategy ensures that permanent unique identification numbers are not frequently exchanged over the air, thus reducing the time in which an attacker can effectively make use of a compromised identification number.

The identifiers include (i) IMSI – This is the permanent user identity which is sent in the clear the first time a user tries to connect to the network. (ii) IMEI – International Mobile Equipment Identity which is the permanent identifier unique to each mobile device. Thus, if a mobile phone is stolen, it is possible to ensure that the network rejects connection requests from this phone even if the SIM card is replaced (iii) M-TMSI – M-temporary TMSI. This temporary identifier is used to identify the UE within the MME. It is assigned by the visited network after encryption. Use of this identifier mitigates against identity-theft security compromise. However, it does not mitigate against subscriber location tracking at the granularity within a cell-site [13] (iv) S-TMSI – This identifier is used for paging the UE. (v) GUTI – Globally unique temporary UE identity. The GUTI is used to enable confidentiality of subscriber identity. GUTI uniquely identifies the MME that allocated the GUTI and identifies the UE within that MME. (vi) C-RNTI – Cell Radio Network Temporary Identifier. It provides a unique and temporary UE identification at the cell level. It is assigned by the network when a UE is associated with a cell [14].

2) LTE End-To-End Security

a) Authentication and Key Agreement

Mutual authentication of the UE and network is a cornerstone of the LTE security framework. The AKA procedure is utilized to achieve this by ensuring that the serving network (SN) authenticates the user's identity and the UE validates the signature of the network. Apart from mutual authentication, AKA is utilized to generate ciphering and integrity keys, which are ultimately used to derive different session keys for encryption and integrity protection.

Three nodes are involved in the authentication procedure – the UE, MME and HSS. The HSS holds subscriber information and is able to verify an authentication request from the UE as well as generate authentication data which it then provides to the MME for processing. The message flow for the UE authentication process starts whenever it tries to attach itself to the EPS. i.e. to register itself with the network.

b) Confidentiality & Integrity of Signaling

The Network Access Control plane includes: (i) RRC signaling between the UE and eNB (ii) NAS Signaling between the UE and MME, and (iii) S1 interface signaling. As

a security feature of the control plane, both RRC and NAS layer signaling are ciphered (encrypted) and integrity protected. Ciphering and integrity protection of RRC signaling is carried out at the PDCP (Packet Data Convergence Protocol) layer while the NAS layer itself is responsible for encryption and integrity protection of NAS level signaling. S1 interface signaling protection is optional. Such protection would not be implemented uniquely for each UE connection. Rather, it would operate over a trusted communication between the eNB and S-GW.

c) User Plane Confidentiality

As a security feature in LTE, user plane data/voice is ciphered between the UE and eNB. Ciphering is optional on the S1-U interface between the eNB and the S-GW. IPSEC-based tunnels can be established between the eNB and S-GW and used to carry the user plane data. No UE-specific security tunnels are established between the eNB and S-GW. As a result, such encryption between the eNB and S-GW is carried out at the IP layer. No integrity protection is provided for the user plane due to performance considerations. The PDCP layer is utilized to facilitate ciphering/deciphering of the user plane for traffic transmitted between the UE and eNB.

VI. Security Issues – 4G Wireless

A. Physical Layer Issues

Both WiMAX and LTE are subject to two key vulnerabilities at the physical layer - Interference and Scrambling attacks [15].

By deliberately inserting man-made interference onto a medium, a communication system can stop functioning due to a high signal-to-noise ratio. There are two types of interference that can be carried out: (i) noise and (ii) multi-carrier [16]. Noise interference can be performed using White Gaussian Noise (WGN). In the case of Multi-carrier interference, the attacker identifies carriers used by the system and injects a very narrowband signal onto those carriers.

Interference attacks can be easily carried out as the equipment and knowledge to carry out such attacks are widely available. Our analysis indicates that interference is easy to detect using radio spectrum monitoring equipments. Using radio-direction-finding tools, the interfering source can be traced. In addition, increasing the power of the source signal and using spreading techniques can increase its resilience against interference. While the possibility of interference is significant, since it is easy to detect and address, we believe its impact on the WiMAX/LTE network and users will be limited.

Scrambling is a form of interference which is activated for short intervals of time. It is targeted against a specific frame or parts of frames. The attacker may target management or control information of a particular user to disrupt service. However, the attacker has to be sophisticated and knowledgeable since specific frames and time slots must be identified for the attack to be successful. As a result, scrambling is difficult to implement successfully. At the same time, we note that it will be difficult to detect and identify an attacker since the signal will be intermittent.

Since most control signals and user data are encrypted in LTE and WiMAX, mounting a scrambling attack on specific control signals or specific user data is very difficult. Therefore we believe that it represents a minimal security concern.

B. WiMAX – MAC-Layer Security Issues

The IEEE 802.16 radio interface standard describes several steps in order for a MS to establish initial access with a Base Station. These steps are (i) Scanning and Synchronization (ii) UL Parameter Acquisition (iii) Initial Ranging and Time Synchronization (iv) Basic Capabilities Negotiation (v) MS Authorization and Key Exchange (vi) Registration with the Serving BS (vii) Connection Establishment. The first five steps involve non-secure traffic. Thus, they are prone to various attacks. Steps 6 and 7 involve secure traffic exchange based on the device authentication standards of WiMAX.

There are various sources of potential vulnerabilities in WiMAX 802.16e [15][17][18][19]. Some of these sources include: (i) The fact that management MAC messages are never encrypted providing adversaries an ability to listen to the traffic and potentially gain access to sensitive information (ii) The fact that some messages are not authenticated (no integrity protection). Typically, a hash based message authentication code (HMAC) is used as digest. However, this is not used for broadcasts and a few other messages. Simple forgery can affect communication between an MS and BS (iii) weakness in authentication and authorization procedures is an enabler for the BS or SS masquerading threat. It is not easy to get the security model correct in a mobile environment due to limited bandwidth and computation resources (iv) Issues with key management such as the size of the TEK identifier and TEK lifetime are considered as potential sources of vulnerabilities for WiMAX security.

Below we present 4 categories of attacks at the MAC layer.

1) Denial of Service

Denial of Service (DoS) attacks are a concern for WiMAX networks. A DoS attack can be initiated via simple flooding, attacking unauthenticated management frames. In one case, the SS/MS authenticates the BS using PKMv2 RSA authentication. In this scenario, the BS has to sign and reply with its public key. Processing of public key encryption and signature is CPU intensive. If flooded with false requests, the BS will be very busy computing and evaluating digital signatures and will be unable to serve any other requests [19].

In a second case an adversary eavesdrops and captures the Authorization Request message from a particular SS to a BS. The attacker then replays the captured message repeatedly. This will burden the BS which will then decline requests from other authentic SS devices.

In a third case, unauthenticated management frames could be maliciously exploited by attackers. Examples include: (i) The Neighbour advertisement message (MOB_NBR-ADV) is not authenticated. The BS broadcasts this message to notify

MSs about the characteristics of neighbouring BSs in order to facilitate handover decision-making. A forged message can announce a wrong or non-existent BS, thus causing denial of service [21][22] (ii) The Fast Power Control (FPC) message is broadcasted by the BS to one or more MSs to adjust their transmit power. This forged message can reduce the power to a minimum, thus causing some MSs to lose connectivity to the BS. It can also set the transmit power to a maximum with the intention of draining batteries [22] (iii) Clock comparison (CLK-CMP) message is a broadcast message and is not authenticated. A forged message can de-synchronize the clock, thus causing denial of service [19] (iv) The Downlink Burst Profile Change Request (DBPC-REQ) message is sent from BS to change the burst profile (modulation, encoding etc.) of the MS when the distance varies between the BS and MS. A man in the middle attack can change a profile, causing complete loss of connectivity between the MS and BS [22]

2) Service Degradation

MAC Management messages are never encrypted and not always authenticated. This can lead to man-in-the-middle attacks causing service degradation as well as DoS attacks. Variants of such attacks are discussed below.

In one example, the Traffic indication message (MOB_TRF-IND) is an unauthenticated broadcast message. It is used by the BS to inform a sleeping MS that there is traffic destined to it. This message can wakeup as many as 32 MSs. A falsely generated message can simultaneously drain the battery of up to 32 MSs [22].

A second example involves the Ranging Request (RNG-REQ) message - the first message sent by an SS seeking to join a network. The message requests transmission timing, power, frequency and burst profile information. An attacker can intercept the message in the middle and change the profile to downgrade the service. Similarly, the Ranging Response (RNG-RSP) message which is sent by a BS in response to the RNG-REQ message carries various profiles. The message is unauthenticated, unencrypted and stateless, thus enabling a man-in-the-middle attacker.

The third example is based on eavesdropping. Management signaling communication is not encrypted and all exchanges between BS and MSs can be listened to. An adversary can create a detailed profile of MSs which may include device capabilities, security settings etc. In particular, monitoring of MAC addresses during ranging can reveal mapping of devices and user equipment [15].

3) Authorization Vulnerability

The authentication/authorization protocol is vulnerable to certain replay attacks. In such a case, the attacker replays an instance of the Authorization Request message sent earlier. The BS then responds with an Authorization Reply message. The BS cannot ignore duplicates since it may be a legitimate duplicate request from the SS due to loss of the previous Authorization Reply message. The BS has to sign and reply with its public key. Processing of public key encryption and signature is CPU intensive. If flooded with replay attacks, the BS will be busy computing and evaluating digital signatures. As a result, it will have little CPU power left to serve any other SS [19] requests.

The attacker can send Authorization Reply messages and gain control of the SS's communication. This man-in-the-middle attack is prevented in 802.16e with mutual authentication support in PKMv2 [17].

In addition the authorization message does not include a digest to facilitate integrity checks and prove that the message has not been modified. This is a potential security deficiency and allows possible replay-attacks.

In another example, the first message in the authorization process is not secured when the SS/MS notifies the BS about its security capabilities (i.e., before negotiating encryption keys). A spoofed message indicating weaker capabilities may downgrade the security strength between a SS and BS. The standard is unclear on the handling of such a scenario [19].

4) Security Issues with Key Management

Key management at the SS has been designed to safeguard it from replay attacks. The SS can determine if a Key Reply message is new or old. This is possible since the old TEK (Traffic Encryption Key) and new TEK are included in the Key Reply message. However, if an attacker replays Key Request messages to the BS, it can trigger frequent exchange of keying materials. This will cause confusion at the SS and exhaust resources at the BS [23].

Another issue arises from the combination of the TEK lifetime and crypto algorithm deficiency. The TEK lifetime can be set to a value ranging between 30 minutes and 7 days. It is known that with the DES-CBC algorithm, security beyond 232 data blocks (each data block is 64 bits) using the same TEK can be compromised. The data may be vulnerable if the TEK lifetime is set to a large value [24].

A third issue involves key management in multicast and broadcast services. 802.16e uses a common group traffic encryption key (GTEK) for traffic encryption/decryption. Each multicast group member must know this key. The transfer of GTEK to all groups is broadcast but encrypted with the shared key encryption key (SKEK). The issue of backward and forward secrecy is not addressed. When a new member receives the current GTEK, it can decrypt all previous messages that were multicast during GTEK's lifetime. In addition, nothing in the protocol prevents the SS after it leaves its group from receiving the next GKEK [23].

C. LTE – MAC Layer Security Issues

One approach to categorizing LTE security issues is to group them as follows [10]: (i) illegal use of user and mobile equipment identities to access network services (ii) user tracking based on the temporary user identifiers, signaling messages etc (iii) illegal access and usage of security procedure keys to access network services (iv) malicious modification of UE parameters (e.g. failure timers, retry timers) to lock out a UE from normal services (v) willful tampering of the eNB system broadcast information (vi)

eavesdropping and illegal modification of IP packet contents (vii) Denial of Service attacks launched on the UE or eNB (viii) data integrity attacks (signaling or user data) using replay. Despite enhancements potential security issues need to be addressed within LTE – categorized below into 4 key types.

1) Location Tracking

Location tracking refers to tracking the UE presence in a particular cell or across multiple cells. Location tracking as such does not pose a direct security threat, but it is a security breach in the network and can be a potential threat. Location tracking is made possible by tracking a combination of the Cell Radio Network Temporary Identifier (C-RNTI) with handover signals or with packet sequence numbers as described below.

The C-RNTI is a unique and temporary UE identifier (UEID) at the cell level. As the C-RNTI is transmitted in clear text, a passive attacker can determine whether the UE using the C-RNTI is still in the same cell or not. During handover, a new CRNTI is assigned to the UE via the Handover Command message. A passive attacker can link the new C-RNTI from the Handover Command message and the old C-RNTI unless the allocation of C-RNTI itself is confidentiality protected. This allows tracking of the UE over multiple cells [14].

If continuous packet sequence numbers are used for the user plane (RLC, PDCP) or control plane (RRC, NAS) packets before and after a handover, then mapping between the old and new C-RNTI's is possible based on the continuity of packet sequence numbers.

2) Bandwidth Stealing

Bandwidth stealing could emerge as a security issue in LTE. In one example, this can be achieved by inserting messages during the DRX period [14]. During a DRX period in the E-UTRAN, a UE is allowed to stay in active mode, but turn off its radio transceiver to save power. During such a DRX period, the UE's context (e.g. C-RNTI) remains active in the eNB. During a long DRX period, the UE is still allowed to transmit packets because the UE may have urgent traffic to send after entering the DRX period. This may create a potential security hole. It is possible for an attacker to inject a C-PDU by using the C-RNTI of a UE during a long DRX period.

In a second example, fake buffer status reports can be utilized. The buffer status report is used as input information for packet scheduling, load balancing, and admission control. Sending false buffer status reports on behalf of another normal UE can change the behaviour of these algorithms. By changing the packet-scheduling behaviour at the eNB, it is possible to carry out a bandwidth stealing attack making the eNB believe that the UE does not have anything to transmit.

3) Security Issues Due to Open Architecture

The 4G LTE network will be an IP network with a large number of devices which are highly mobile and dynamic with activity periods ranging from a few seconds to hours. The types of end-devices will be very diverse and will include a heterogeneous range of end-users. Additionally, a broad range of automated devices are emerging which operate without human interaction. Such devices take advantage of the ubiquity of wireless network coverage and include for example sensors, alarms, presence indicators and remote cameras.

Diversity in device types and security levels coupled with the open architecture of an IP-based LTE network will result in greater numbers of security threats than seen in 3G networks. At present, hand-held mobile devices (mainly cellular phones) are the most wide spread users of wireless networks. Such devices have typically been proprietary in their design and makeup. While there is initial evidence of malicious activity in cellular networks, large scale infection of cellular smart phones has not yet occurred. As one example, in [25], the authors carried out a study from a 3G network in Europe. They studied data traffic from laptops running Microsoft windows in an Austrian service provider. The study revealed that a large fraction (50%) of uplink packets on laptops with UMTS cards were TCP SYN packets directed to the ports TCP:135 and TCP:445 from infected user devices.

4) Denial-of-Service (DoS) Attacks

In LTE networks, there may be two possible ways to carry out a DoS. The first type of DoS attack would be against a specific UE. A malicious radio listener can use the resource scheduling information along with the C-RNTI to send an uplink control signal at the scheduled time, thus causing a conflict at the eNodeB and service problems for the real UE.

Newly arriving UEs are susceptible to a second type of DoS attack. In LTE, the UE is allowed to stay in active mode, but turn off its radio transceiver to save power consumption. This is achieved via the DRX (Discontinuous reception) period. During a long DRX period, the UE is still allowed to transmit packets because the UE may have urgent traffic to send. However, this can create a potential security hole. For example, attackers can inject C-PDU packets during the DRX period to cause DOS attacks against newly arriving UEs.

A third type of DoS attack can be based on the buffer status reports used by an eNB for packet scheduling, load balancing, and admission control. Attackers can send reports impersonating a real UE. If the impersonator sends buffer status reports which report more data to send than are actually buffered by the real UE, this will cause a change in the behaviour of admission control algorithms [14]. If the eNB sees many such fake buffer status reports from various UEs, it may believe that there is a heavy load in this cell. Consequently, the eNB may not accept newly arrived UEs.

D. Security Issues at the Higher Layers

It is expected that a range of security risks will emerge in 4G wireless due to a number of factors including: (i) departure from proprietary operating systems for hand held devices to open and standardized operating systems and (ii) open nature of the network architecture and protocols (IP-based). With this move to open protocols and standards, 4G wireless networks are now susceptible to computer attack techniques present on

the Internet. Such networks will be increasingly vulnerable to a range of security attacks including for example Malware, Trojans and Viruses [26]. Apart from end-user equipment posing traditional security risks, it is expected that new trends such as SPIT (SPAM for VoIP) will also become a security concern in 4G LTE and WiMAX. Other VoIP-related security risks are also possible such as SIP registration hijacking where the IP address of the hijacker is written into the packet header, thus, overwriting the correct IP address [18].

VII. CONCLUSION

This study of security issues in 4G networks has revealed that both WiMAX and LTE security architectures are at advanced stage of specification. This study focused primarily on MAC layer vulnerabilities for WiMAX and LTE. Both standards also have some physical layer vulnerabilities to interference and scrambling techniques. At the MAC layer, WiMAX is susceptible to DoS attacks, eavesdropping, replay attack, service degradation, and vulnerabilities due to faulty key management. Some of these vulnerabilities have been addressed in subsequent versions of the standard (e.g., 802.16e and 802.16m). LTE also has a set of potential vulnerabilities at the MAC layer. Examples of specific vulnerabilities include: illegal use of user and mobile equipment, location tracking, DoS attacks and data integrity attacks.

The robustness and effectiveness of end-to-end security approaches in WiMAX and LTE will become clear only after deployment. While both standards improve on their predecessors, there is clearly still work to be done. We believe there is a strong need for continued study on 4G security issues and development of appropriate counter measures. To date, majority of the research work has focused on studies or preliminary simulations. We suggest that there is a critical need to augment this initial research with emulation and test-bed related studies which will likely reveal further issues and challenges to be addressed.

ACKNOWLEDGMENT

We are grateful to Randy Sutton and Stephen Jaworski for numerous discussions on issues related to mobile wireless networks including 2G, 3G and 4G.

REFERENCES

[1] ITU-R M.1645: Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000

[2] "UMTS LTE Network Architecture", Technical specification TS 23.002, version 8.4.0 (Release 8), 3GPP

[3] Zhang and Y. Fang, "*Security Analysis and Enhancements of 3GPP Authentication and Key Agreement Protocol*", IEEE Transactions on Wireless Communications, Vo. 4, No. 2, March 2005

[4] M. Shin, A. Mishra, J. Ma, and W. Arbaugh, "*Wireless Network Security and Interworking*", The Proceedings of IEEE on Cryptography and Security 2005

[5] S Putz and R. Schmitz, "*Secure Interoperation between 2G and 3G Mobile Radio Networks",* First Int Conference on 3G Mobile Communication Technologies, pp. 28-32, March 2000.

[6] G. Horn and P. Howard, "*An Introduction to the Security Features of 3GPP and Third Generation Mobile Communications Systems*", IEEE VTS 51st Vehicular Technology Conf, May 2000, Tokyo

[7] M Bargh et al, "*UMTS-AKA and EAP-AKA Inter-working for Fast Handovers in All-IP Networks*", Workshop on Security and Privacy in 4G Networks, IEEE Globecom, Nov 2007 Washington DC.

[8] V Niemi and M. Blommaert, "*3GPP Security Hot Topics: LTE/SAE and Home (e)NB*", 4th ETSI Security Workshop, Jan 2009

[9] 3GPP TS 33.401 V8.2.1, 3rd Generation Partnership Project; "3GPP System Architecture Evolution (SAE): Security Architecture, (Release 8)", Dec 2008.

[10] C.B. Sankaran, "*Network Access Security in Next Generation 3GPP Systems: A Tutorial*", IEEE Communications Magazine, Feb 2009.

[11] D Tonesi, A. Tortelli and L. Salgarelli, "*Security Overheads for Signalling in Beyond-3G Networks*", IEEE Int Workshop on Digital Communication, TIWDC, Italy, Sept. 2007.

[12] A. R.Prasad et al, "*Mobility and Key Management in SAE/LTE*," IEEE Int Workshop on Digital Communication, Italy, Sept. 2007

[13] B. Ravishankar and M. Harishankar, "*Roaming Issues in 3GPP Security Architecture and Solution using UMM Architecture*", 2nd Conf on Mobile Ubiquitous Computing Systems, Services and Technologies.

[14] D. Forsberg, et al, "*Enhancing security and privacy in 3GPP E-UTRAN Radio Interface*", The 18th IEEE International Symposium on PIMRC, Athens, Sept 2007.

[15] M. Barbeau, "*Wimax/802.16 threat analysis*", Proceedings of the 1st ACM international conference on Quality of Service & security in wireless and mobile networks. New York, 2005

[16] M. Husso, "*Performance Analysis of a WimAX System under Jamming*", MSc thesis, Dept of Electrical and Communication Eng, Helsinki University of Technology, Finland, Dec 2006.

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

[18] Y. Park and T. Park, "*A survey of Security Threats on 4G Networks*", IEEE Globecom Workshop on Security and Privacy in 4G Networks, November 2007, Washington, DC.

[19] P. Rengaraju et al, "*Analysis on Mobile WiMAX Security*", IEEE TIC-STH Conf - Symposium on Information Assurance, Sept 2009, Toronto

[20] L. Maccari; M. Paoli; R. Fantacci, "*Security Analysis of IEEE 802.16*", IEEE ICC '07, June 2007, Glasgow, Scotland.

[21] T. Han, N. Zhang, K. Liu, B. Tang and Y. Liu, "*Analysis of mobile WiMAX security: Vulnerabilities and solutions*", 5th IEEE Int Conf on Mobile Ad Hoc and Sensor Systems, Sept 2008, Atlanta

[22] A. Deininger et al, "*Security Vulnerabilities and Solutions in Mobile WiMAX",* Int Journal of Computer Science and Network Security, Vol. 7 # 11, Nov 2007

[23] C. Huang and J. Chang, "*Responding to Security Issues in WiMAX Networks*", IT Professional, Vol 10, Issue 5, Sept-Oct 2008.

[24] "WiMAX: Standards and Security", ed: S. Ahson and M. Ilyas, CRC Press, 2008

[25] F. Ricciato, P. Svoboda, et al, "*On the Impact of Unwanted traffic onto a 3G network*", FTW. Tech Report, Feb 2006.

[26] JF Beaumont and G. Doucet, "*Threats and Vulnerabilities of Next Generation Satellite Personal Communications Systems: A Defence Perspective*", IEEE Globecom Workshop on Security, and Privacy in 4G Networks, Nov 2007, Washington, DC