3GPP TSG SA WG3 Security — S3#25 **S3-020524**
8 - 11 October 2002
Munich, Germany

Source: Ericsson

Title: 3G-WLAN - Security Evaluation and Countermeasures Proposal

Document for: Discussion and approval

Date: 17-09-02

1 Introduction

This document describes potential threats and attacks, and proposes countermeasures for a 3G-WLAN interworking solution. The purpose is to help 3GPP identify security requirements for the 3G-WLAN system, and choose suitable security mechanisms fulfilling those requirements.

2 Security for Public WLAN Access

These questions related to security in the 3G-WLAN architecture must are addressed:

- What needs to be protected? I.e., what are the assets, and to whom are they valuable?
- What trust relations can be assumed? I.e., who can trust whom, and to what degree? The Trust Model is described in a separate paper.
- What are possible attacks against the assets, how can they be performed, and how what could be done to detect/prevent them?

2.1 Assets

This section describes different types of assets that are valuable to the parties involved. Threats to these assets are also identified.

2.1.1 Network access

The business model of public access hinges on the operator's ability to correctly charge for network access. Operators must consider that dishonest users may attempt to gain free access. Different charging models lead to different threats.

Flat rate charging

If the user pays a monthly fee, and then can access the network as much as he wants, the user might be less careful with his credentials; it does not really matter to the user if someone else is using his account, as far as his expenses are concerned. Of course, he still would not like someone doing anything illegal, using his identity. From the operator's perspective, this type of double usage can be viewed as a lost customer.

Another possibility is to charge the user based on shorter periods, e.g. on a per hour basis. This may however be thought of as a flat-rate model.

Volume-based charging

The user could pay a certain amount of money for a fixed number of bytes transferred to or from his device. With this model both the user and the operator will be careful with authentication and accounting. This does, however, come with additional costs for implementing accounting for the operator. It also gives the problem of handling flooding attacks against user devices, initiated from within Internet.

2.1.2 Cellular Operator's Network

From the operator's perspective, it is important that the security of its physical networks and its ordinary cellular subscribers is not jeopardised. This can only be accomplished if the added WLAN-functionality has a security level that is comparable to the one in the cellular access network.

2.1.3 Wireless Access Provider's Network

The WLAN users must be able to access the network at any time. This implies that the operator needs to be able to detect and to the extent possible protect against Denial of Service (DoS) attacks, both from the Internet, and through the access network.

2.1.4 User Data

For the WLAN users, the integrity and confidentiality of the transmitted data is of importance. That is, the user must feel confident that these properties are guaranteed. In addition, the user must be sure that no one else can access the network using his credentials to perform illegal activities or simply gain "free access" on the user's bill.

This evaluation focuses on security for network access, hence the protection of user data residing on the user's device (such as corporate documents stored on a laptop, for example) is not considered.

Once the user is connected to the IP-network, he is open to all threats that he would be while connected by any other means. If a company, for instance, would like to have a secure tunnel from the user's machine into their network, they should follow the same procedure as they would if the user performed the same task from his home, using an ISP.

2.2 Attacks

This section identifies attacks that are applicable in a typical WLAN Access Network and discusses countermeasures against these.

2.2.1 Wireless "Boot-strap" Attacks

To perform any type of attacks in the WRAN, the attacker needs access to the network in some way. For ordinary wired networks, an attacker needs to somehow hook up to the wires to get access. The WRAN itself is partially a wired network, and an attacker may hook up to that part of the network as well. The nature of wireless networks however, opens up for easier solutions (both for the normal users and for attackers). This section describes some ways a malicious person can get access to the WRAN.

For certain types of attacks, the perpetrator does not need to "be a part" of the network. Examples are some types of ARP attacks (see below), and certain DoS attacks, e.g., setting up a radio jammer in a hotspot. It should, of course, also be noted that an easy way of getting access to the WRAN is to simply become legitimate subscriber.

Rogue AP attack

In a rogue AP attack, the attacker employs an AP (masqueraded as a legitimate AP in a given hotspot) connected to an MS, as depicted in Figure 1. Based on signal strength, an unsuspecting MS may connect to the rogue AP and start to perform authentication. Since no messages can be integrity protected before authentication, the attacker substitutes the MAC/IP address-pair of his own MS and relays the authentication messages to a legitimate AP. In this way, the authentication procedure binds the MAC/IP address-pair of the attacking MS to the credentials of the legitimate user. As a consequence, the attacker gains access to anything the legitimate user would, while the legitimate user is denied access. This attack is only applicable if the authentication does not result in encryption/integrity keys to protect the session.

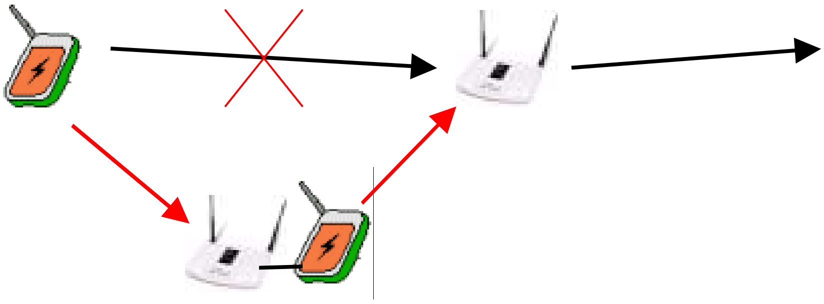

Figure 1. Rogue AP used for man-in-the-middle attacks

The same equipment may also be used to redirect a user's traffic to a completely different network. That is, the attacker may trick the user into believing that he is accessing the given WRAN, when he in fact is connected to a network of the attacker's choice.

Impersonation attack

Another way to gain access for an attacker is to eavesdrop on the traffic around an AP. Since the MAC and IP addresses are sent in the clear (they are not encrypted) the attacker can record these. When the attacker knows the MAC/IP address-pair of a user currently connected, he can set his own addresses to the same values. When this is done, the attacker can send and receive data through the AP. However, it is unclear whether two users accessing the network through the same AP will disturb each other to such a degree that the usability is non-existent. But, if the access network does not keep track of the location of its users, the attacker may move into the coverage of another AP and continue the session from there. Furthermore, if the legitimate user leaves the hotspot, the attacker may continue the session afterwards. Again, a prerequisite for the attack is that there is no encryption or message integrity checks of the transmitted data during the session.

Countermeasures

Since a new AP that is introduced in a hotspot (by an attacker) will create radio disturbance, it can be detected. The WLAN access provider's AP:s should be enabled to detect such disturbances and notify the provider/operator when any anomalies are discovered. Ideally, the WLAN provider would like to remove/disable rogue AP:s, but that may not be a legally viable option given that 802.11b WLAN:s operate in unregulated ISM band.

Integrity and/or encryption keys should be derived in a secure manner during the authentication procedure, and these should be used to protect the traffic of the session. An attacker that uses the "relay-attack", where the MAC/IP addresses of a legitimate user are bound to a rogue MS, would then not be able to manipulate or sniff any of the data that is passing through it; thus, making the attack fairly useless.

To prevent someone from masquerading as another user using his IP/MAC address pair, and keep on using it after the legitimate user left the area, is to request logout, or performing re-authentication at regular intervals.

As mentioned in the previous paragraph, all traffic between the user and the AP should be integrity protected. This would also protect against an attacker spoofing MAC/IP addresses to gain access on a legitimate user's account. Although less effective, periodic re-authentication can be used to limit impersonation attacks to the duration when legitimate users are actually present at a hotspot.

2.2.2 User Data Attacks

Once an attacker has successfully become a man-in-the-middle, he could present a fake login-screen to the user. This screen could ask the user to re-authenticate (because of some fictious error) or pose as the original login-screen (if the user hasn't yet authenticated himself). From this login screen, the attacker can then steal the user's credentials.   This attack relies on the general inability of users to verify certificates.

Another possibility for a man-in-the-middle attacker is to divert the user's traffic to a network other than the WRAN the user intended to use.

An attacker can easily eavesdrop the traffic between a user and an AP. The only equipment needed to do this is a laptop with a WLAN interface.

If the volume based charging model is applied, an attacker could flood a user with garbage packets, just to increase the user's bill. This is effective if the attacker resides somewhere on the Internet with a flat rate charging model, or if the attacker has infected other users' machines with "bot"-software. Bot is short for robot, and refers to software that "lives on its own". The bot could for instance listen for connections on a certain port, and when receiving a command from the attacker on that port, it starts flooding a given IP address with packets. Various distributed denial of service (DDoS) tools using such bots are known and available in the hacker world.

Countermeasures

Since the attacker needs to be a man-in-the-middle to perform the fake login-screen attacks, one possibility to avoid this is to use the same countermeasures as in Section 2.2.1. That is, detect false AP:s and remove them if possible.

Also, integrity protection of the data from the MS to the AP would make it impossible for an adversary to inject packets in the data stream, and confidentiality protection would prevent eavesdropping.

If a login-screen is used, it should be accompanied by some sort of certificate. The users must be educated on how to verify them. The verification must be simple enough for an everyday user to carefully perform it every time.

To counter flooding/garbage attacks from the Internet, the WLAN site could be protected by an intelligent firewall, i.e. a firewall that can distinguish legitimate traffic from bogus traffic with high probability.

2.2.3 IP Network Attacks

The attacks applicable to general IP networks are, of course, also applicable in the WRAN. The fact that there are more ways of getting access to a wireless network than to a wired one makes these attacks even more serious.

This section will describe some of the most common attacks possible in IP networks. An important class of attacks are "service spoofing" attacks, where the attacker impersonates one or several services/servers in the network, e.g., a DNS server or a DHCP server. Another set of attacks uses fake configuration/control messages (such as ARP or ICMP messages) to redirect a user's traffic. Note that the above include only the best-known and most serious attacks. Given the rich (and always expanding) set of protocols run over IP, all possible attacks could not be accounted for.

ARP

The Address Resolution Protocol (ARP) is used to map IP addresses to MAC addresses. When the user's MS wishes to retrieve the MAC address of a machine, it broadcasts an ARP message, basically asking "who has IP xxx.xxx.xxx.xxx?". The machine that has that IP address should then respond to the calling host. When the host receives the response, he updates his ARP-cache with the given IP-MAC address mapping. A problem with ARP is that it is stateless. That is, a host that receives a response message will update its ARP-cache whether he sent a query message earlier or not. This feature can be exploited by sending fake ARP-replies to any host and in this way force his IP addresses to map to arbitrary MAC addresses. This is known as "ARP-poisoning".

By tricking two hosts into believing that the attacker's MS is "the other host", an attacker can implement man-in-the-middle attacks between the two.

It should be noted that an attacker does not need to be part of the network to be able to perform these attacks. By just being present in the hotspot, he can send and receive ARP-messages. Using standard equipment, the user may be part of the network on layer 2, but does not need to be part on layer 3. Using slightly more sophisticated equipment, he does not even have to be part of the network on layer 2.

DHCP

On most access networks, the Dynamic Host Configuration Protocol (DHCP) is used to assign IP-addresses to visiting clients. The DHCP server will respond to client requests with an IP address from its pool. As is the case for DNS queries (see next paragraph), any attacker able to respond before the legitimate DHCP server, can respond with a bogus address of his choice.

DNS

When a user needs the IP address corresponding to a domain name, he broadcasts a Domain Name Service (DNS) query message. Any user on the same network segment will see these messages. An attacker that responds quicker than the DNS can then send a fake reply to the query, and can effectively tell the user that the queried domain name corresponds to any IP address he chooses. When the reply from the actual DNS server arrives to the user, it will be discarded, since the MS has already received a reply.

ICMP

The Internet Control Message Protocol (ICMP) is designed to aid hosts in routing their traffic. There are several possible attacks based on ICMP. For instance, the so-called Smurf-attack is done in the following way: the attacker broadcasts an ICMP echo request (ping message) message on the network, where the return-to address is changed for the victim's address. All hosts in the segment will respond to the ping, and the victim is flooded by ICMP packets (this is the analogue of the Fraggle-attack, which uses UDP echo packets in the same manner). Other possible attacks include sending an ICMP redirect (route update) message to the victim, which will change the victim's routing table.

Countermeasures

These attacks are possible whenever the MS:s are located on the same network segment. As can be seen above, they all depend on the attacker being able to see the other users' traffic. This can all be avoided by disallowing unwanted communication between MS:s directly through the AP:s, and by forcing all traffic from the AP:s through another node sitting further within the WLAN access network.

This new functional entity –known as Access Server Node (ASN), Network Access Server (NAS), WRAN Support Node (WSN) or other similar names– is not part of the 3G-WLAN architecture today, and it's been subject to debate in IEEE lately.

The Access Server Node connects several AP:s and handles the signalling to the AAA server in the Home Network. The ASN forwards the traffic from the WRAN either directly to the Internet or to a node in the operator's service network. In that case the service network may contain an access server that relays the traffic to the Internet, possibly performing packet filtering, accounting, etc.

Typically, the ASN would implement a traffic filter that allows only packets from authenticated users to pass and blocks everything else. To distinguish between authenticated users and others, the MAC/IP address pair of the packets is usually checked. That is, once a user is authenticated, the ASN opens a hole in the packet filter for the corresponding MAC/IP address pair

It should be noted that this does not prevent two MS:s from talking to each other. They could use ad-hoc mode, or they could send packets to each other through the network as well. The important thing is to not let MS:s spoof services by preventing non-legitimate MS:s to answer service queries.

2.2.4 Attacks on the WRAN and Countermeasures

Attacks on Layer 2 Mobility Messages

If a user moves between the coverage area of two AP:s within the same WRAN, signalling on layer 2 is done to perform the handover. There is no standard protocol for this signalling yet, but the IEEE is working on a solution called Inter-AP Protocol (IAPP). If no precautions are made, an attacker may be able to send mobility messages, telling the WRAN that a user has switched to another AP. This could result in a simple DoS attack. If the attacker having spoofed the MAC/IP addresses of a user and simulates a move into the coverage of another AP, he will get better performance in his stolen access due to fewer collisions on the air interface. If the network is configured to drop the user at the old AP when he moves, the legitimate user will lose his access, while the attacker maintains his.

To counter this, the layer 2 signalling should be disallowed, or should at least be integrity protected.

Attacks on Switches

To counter some of the general IP network attacks mentioned above, the network may use a switch to make sure that all traffic is not broadcast to all AP:s (and over their air interface). Some switches have the vulnerability that under heavy load they start acting like hubs. That is, they start to broadcast traffic. Other switches stop receiving traffic on the overloaded ports altogether. An attacker can easily overload a switch with basically any type of traffic.

Attacks on AP:s

DoS attacks on the AP:s can be difficult to detect, and hence to counter. There is always the question of how to distinguish between a DoS attack and just a high load of legitimate traffic.

2.2.5 Malicious Software in the User's machine

Open platform terminals may be infected by viruses or other malicious software. The software operates without the knowledge of the user on his terminal, and can be used for different types of attacks:

- If the user has credentials stored on a smart card connected to his terminal, a trojan residing in the terminal can make fake requests to the smart card and send challenge-response results to another MS. The owner of the latter MS can then get access with the stolen credentials.
  Note that this attack is performed inside the terminal, and it is independent of the external link between the terminal and the smartcard reader, which can be secured or assumed to be physically secure.

- Trojans may perform all the usual activities: monitor the user's keyboard or sensitive data, and forward the information to another machine.

- Malicious software can be used to perform Distributed DoS (DDoS) attacks. That is, several instantiations of the software (residing on different hosts) synchronise and start a DoS attack simultaneously against a target.

- Malicious software could be trying to connect to different WLAN:s, just to annoy the user.

Alternatively, the (U)SIMin the cellular phone can be used remotely from the WLAN client through a serial, infrared, or Bluetooth connection; in a sense use the phone as a smartcard reader.

As the terminal must access the (U)SIM in the phone, the link in between must be secure. Both cable and IR can be assumed physically secure, and Bluetooth will depend highly on the current Bluetooth security mechanism (which can be assumed secure).

Cellular operator may be wary of using the same long-term secret for both cellular authentication and WLAN access. A separate USIM application for WLAN access could be added to the same physical card. If this is still not acceptable then a separate WLAN module could be issued, but that is likely to require an external smartcard reader and is also less attractive from a reuse perspective.

Using the cellular phone as an intelligent smartcard reader might have an advantage over standalone readers, since the phone could implement extra checking to restrict access and possibly even report to the operator. Such extensions, however, would require an upgrade of current phone software.

Countermeasures

Unfortunately, most mainstream operating systems used on laptops/PDAs today lack support for mandatory access control and/or trusted path mechanisms to smart card readers; in other words, they offer no means to protect against this credential stealing attack.

As the real threat is not simply the trojan's access to credentials, but its assumed ability to export the credentials for use elsewhere, firewall technology may be deployed to monitor suspicious outgoing traffic. However, automatic detection may prove difficult, as it is not immediately clear how to distinguish between legitimate and malicious packets. On the other hand, just like in virus protection or intrusion detection solutions "signature" patterns of a certain type of trojan may be used to identify and remove the culprit.

The protection against malicious software lies very much in the hands of the user himself. The user should be informed of the hazards with downloading untrusted software, be encouraged to use a personal firewall, and regularly update anti-virus protection software.

3 Countermeasures proposal

In this section we turn the problem around, and examine the impact of the different countermeasures on system security. Instead of the attack-countermeasures treatment in section 2, here we look at each countermeasure to see which attacks it prevents (or detects), and try to estimate the resources required to implement it. The results are summarised in Table 1 at the end of this section.

3.1 Firewalls and Packet Filtering

In a most deployment scenarios there will possibly be a firewall on the Internet side to implement access control from the WRAN and protect the nodes at the WLAN site. In addition to preventing attacks on the service nodes, such firewalls can and should be used to also protect WLAN users from malicious traffic. Using stateful firewalls with application-level gateways (ALG:s), most flooding and / or garbage attacks can be detected and effectively avoided. This is essential for volume-based charging. Intelligent packet filters can also recognise and drop packets that match patterns for some of the well-known IP attacks, thereby limiting such attacks to the WRAN where local protection mechanisms (see VLAN:s in the next section) can be applied.

Since firewalls are already a standard part of most WLAN access sites, implementing the above countermeasures requires no extra hardware. As most modern firewalls implement ALG:s, the main cost derives from proper configuration and maintenance.

3.2 WRAN Protection

As described in section 2.2.4, to protect the wired network behind the AP:s from hackers, and to hide the site structure from competitors, all incoming traffic from the wireless interface of the AP:s should be forced directly to an Access Server Node (ASN). This same countermeasure also limits uncontrolled mobile to mobile communication to direct air-link connections, which helps to ensure that IP attacks cannot be launched using the WLAN provider's own infrastructure. Note that rogue AP attacks are still possible.

This countermeasure can be implemented using Virtual LAN (VLAN) enabled switches to force traffic to the desired paths only. Since most site designs already include switching equipment, the only task is to pick ones with VLAN support. The associated costs are relatively minor. The VLAN, of course, have to be configured properly. Some additional configuration on the AP:s may also be necessary.

3.3 Fraud Information Gathering and Intrusion Detection Systems

Although a Fraud Information Gathering System (FIGS) or an Intrusion Detection System (IDS) cannot directly prevent attacks or fraudulent behaviour, by detecting such activity they can activate emergency measures, which may in turn stop an attack or at least contain the damage.

Typical uses of FIGS:s include consistency checks (e.g. the same user logged in twice, or the same account used in geographically distant locations within a short time), as well as keeping user profiles and detecting anomalies from "normal" (previously tracked) behaviour.

IDS:s usually scan otherwise legitimate-looking traffic (that may have passed a firewall) for abnormal traffic patterns and/or well-known attack signatures. They can implement rules to disable users/ports in a firewall. They are also ideal to detect DoS attacks, and activate DoS-robust modes of operation on certain servers. (Such modes are not normally used due to an undesirable performance loss, unwarranted under normal conditions.)

Since most operators already employ a FIGS of some sort and IDS:s are often bundled by firewall vendors, the costs involved can be minimal. These systems are particularly important, as they can be easily adapted to detect attacks not yet known today.

3.4 Integrity Protection

Integrity protection between mutually authenticated parties (UE-AP, UE-ASN, or UE-VPN gateway) protects against masquerading and man-in-the-middle attacks as well as flooding/garbage attacks from the wireless network.

Unfortunately, attacks cannot be prevented during the authentication phase. However, the authentication procedure must be robust enough to complete without any protection under normal circumstances and indicate failure under attack. It must also result in keying material for the subsequent integrity protection.

Mutual authentication and integrity protection would be best implemented on the link-layer. Unfortunately, the current 802.11 standard does not provide adequate protection. The trends in standardisation indicate that link-layer protection will be the long-term solution, but for immediate deployment with off-the-shelf components this cannot be used. In addition, the standard is mainly concerned with security between the MS and the AP, and not between the MS and the ASN.

Using IKE and IPsec can provide the necessary protection at the IP-layer. This requires software and configuration in both the MS:s and the AP or ASN. Fortunately, most modern mainstream operating systems support IPsec either natively, or as a third-party add-on. Concurrent use of IPsec on the client (e.g. for corporate access) may cause unforeseen problems. In addition, running IPsec in the AP:s or ASN:s can potentially harm performance.

Countermeasure	Estimated resources	Protection
Firewalls with ALG:s	minimal/configuration	prevents/reduces external IP attacks, flooding/garbage attacks
VLAN-enabled switches	minimal/configuration	prevents WRAN attacks; limits IP attacks to direct air-link
FIGS / IDS	variable (may need extra HW / SW, if not already deployed)	detects attacks and / or fraudulent behaviour; triggers emergency response
Integrity protection (with mutual authentication)	link-layer: reasonable (long-term)needs new equipment and/or new standardsIP-layer: reasonable (immediate) SW support widely available, server capacity may suffer	degrades masquerading and man-in-the-middle attacks to DoS on user clients; makes IP attacks ineffective; reduces the effect of DoS on servers

Table 1 Countermeasures

4 Proposal

It is proposed that the Security Evaluation and Countermeasures proposal described in sections 2 and 3 are agreed by SA3 and incorporated to Annex C of Technical Specification 33.cde "WLAN Interworking Security".