Mobile Network Security

Availability Risks in Mobile Networks

Author: Yaniv Balmas, ERT Lab Security Researcher

November 2013

Table of Contents

Executive Summary

Mobile networks have left the “walled garden”. A privileged, closed and isolated ecosystem which is under the full control of mobile carriers, used proprietary protocols and has minimal security risks due to restricted user access. With the introduction of 4G, Long Term Evolution (LTE) networks, and the IEEE standardization of mobile networks, the secure, “walled garden” days are over. Mobile networks are becoming very similar to common IP-based networks. However, while organizations have years of experience and knowledge in defending against cyber threats in common IP networks, they are years behind in terms of accumulated knowledge in defending mobile networks.

Key Findings

- Migration has not been met with equal funding – providers have not invested in security at the same rate as capacity.
- Mobile service providers must adapt to upcoming changes in the threat landscape, and be prepared to ensure network availability.
- Since new threats may have potential for catastrophe, the mobile network providers should adapt security programs and procedures to withstand those threats, while assuring the same SLA and preventing any major service outages.

Key Technical Recommendations

As with any other network security program, there is no magic solution that will eliminate all threats. The most effective and robust solution should be the product of a careful risk assessment process, that analyzes all business critical “weak spots” and implements the proper technical and procedural solutions to compound them.

The various technical solutions involved in such a solution include:

- **Deep Packet Inspection (DPI)** – deep packet inspection solutions that were designed to be implemented on mobile networks, on network gateways, in the access networks, and in the core network itself.

- **Behavioral Based Mitigation** – since many threats are zero-day attacks or use legitimate transactions to misuse resources, the only way they could be detected and stopped is by behavioral analysis techniques. These techniques will allow normal network traffic to pass through while mitigating abnormal traffic patterns.

- **VOIP Protections** – solutions that will protect VOIP signaling infrastructure from being exploited from both internal and external sources.

- **DNS Protections** – specific DNS oriented protections that will be able to detect and effectively block any massive abuse of the various DNS services found in the mobile networks.

- **Signaling Proxies** – that will handle all signaling traffic from a single point and make sure the network will not be overwhelmed from the signaling traffic rates.

Introduction

Like any data communication network, mobile networks contain a range of security threats. Though some threats are easy to identify and mitigate, others are illusive, due to the unique structure and complexity of mobile networks.

The transformation to fully IP-based mobile networks involves a transition period, during which existing security vulnerabilities will be exposed to a substantially larger audience. Attackers can easily generate attacks targeting mobile endpoints, overlapping network services between mobile and other networks (such as DNS), and even the core network itself. A full-blown attack on a mobile network has the potential for catastrophic results that affect multiple audiences. In today's hyper connected world it could be viewed as a national infrastructure attack. End user privacy may be breached; mobile network operators can lose considerable revenues; and confidential government data may be exposed. We have already seen several attack cases that were near such catastrophe.

This research paper examines mobile network technologies and provides insight into the specific threat types that may affect mobile network availability today and in the future.

Mobile Networks Overview

To fully appreciate the security risks in mobile networks, one must first understand the basic design and key components that consist of a mobile network.

Network Design

All mobile networks, regardless of their generation, contain the same basic network elements.

- **Mobile Devices** – communicate through radio channels, and exchange voice and data traffic between the carrier network and the user.

- **Radio Access Network** – is the part of the mobile network which connects subscribers to their service provider. An access network typically includes mobile towers and serves as the gateway between radio networks and the core network.

- **Backhaul** – are the physical connections and networks used to carry data between the Radio Access Network and the Core Network.

- **Core Network** – holds network logic and is in charge of creating and maintaining the connection between the mobile devices and the external service networks. It transfers the user data and the control data, authenticating users and devices, calculating billing, and enforcing quality of service, etc.

- **External Service Networks** – contain the end user services and may include connection to PSTN or to VoIP networks, internet browsing services, interconnection to other providers, enterprise specific networks and many other services.

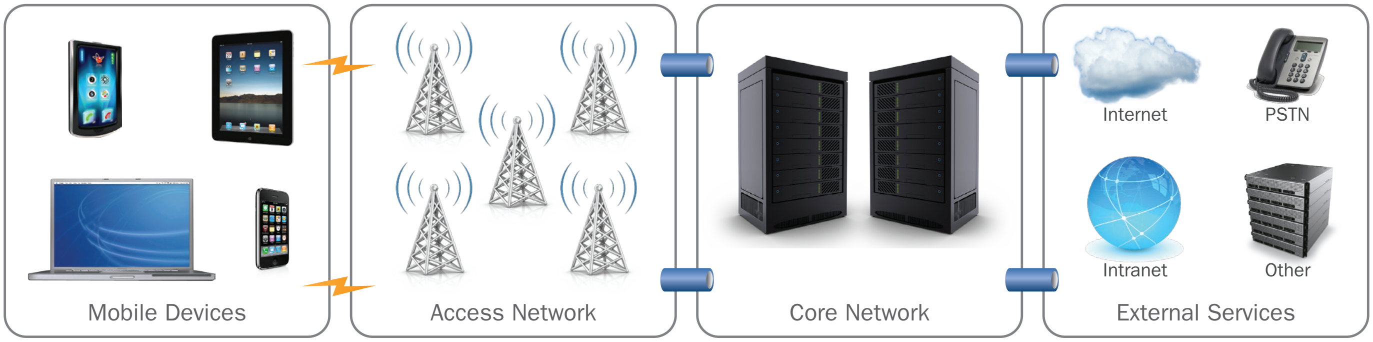

Figure 1 - Basic Mobile Network Elements

Potential Mobile Network Entry Points

Each of the mobile network components may provide attackers with opportunities for gaining entry and abusing the network. The following section reviews the various components and explains the possible ways and resources required for breaking in.

Mobile Devices

Mobile Devices - cell phones, smart phones, tablets, laptops or any other device that provides network connectivity presents the most obvious entry point into the mobile network.

Most modern mobile devices include two key data communication components:

- **Application Processor** – controls the UI components (i.e. screen, keypad, Microphone) and runs the entire software stack that interacts with the user including applications. The application processor usually runs a generic operating system, such as iOS, Android, Windows Mobile, etc.
- **Communication Processor** – runs a cellular protocol stack on top of a real-time operating system (such as Nucleos or Threadx). The PM and RF chips directly connected to the communication processor are responsible for power management and the conversion of baseband to radio frequencies.

  The communication processor application software is usually produced by the mobile device vendor (such as Apple's iPhone) or bought from a third party vendor which distributes it as binary only.

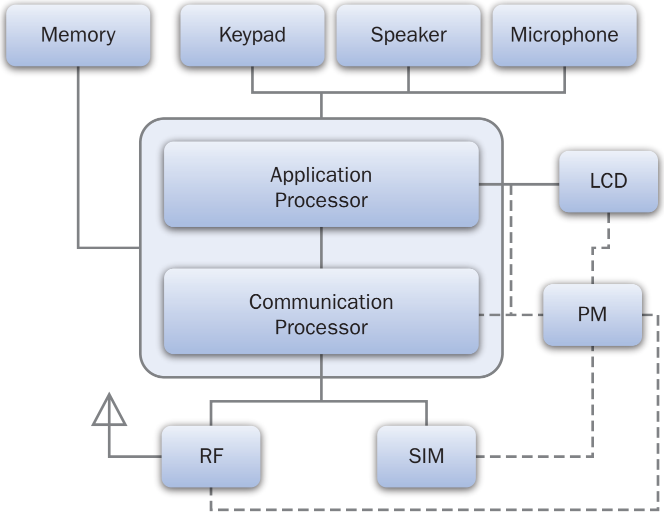

Figure 2 - Generic mobile phone architecture

It is relatively easy for an attacker to gain control of the application layer in an existing off-the-shelf mobile phone. This can be done through insertion of mobile malware, or by using social engineering techniques in order to convince users to perform some actions/commands on their mobile device.

However, this method has its limitations. In many cases of modern mobile devices, gaining control over the application layer will not allow an attacker to control any radio signaling and the attacker will not be able to interfere with the actual inner workings of the mobile network communication.

In order to achieve such a level of control, without owning enterprise grade equipment, the attacker should either exploit vulnerability in the communication layer implementation or construct a low-cost baseband controller. In the last few years several examples and proof of concept projects demonstrated that both cases are quite realistic. [1] [2] [3]

Backhaul

The backhaul can provide an attacker with access to all the control and data traffic sent between mobile devices located in a specific coverage area.

In some mobile network architectures, base stations act as a termination point for encryption protocols. The data can be relatively easy to read or modified and provides the attacker with access to both signaling and user communications going to and from the compromised base station.

While gaining backhaul access does present several challenges to an attacker, compromising a backhaul connection will only provide control of a limited number of mobile devices located in the same geographical location of the point of entry. In order to achieve backhaul access, the attacker may be required to physically break into a base station or, find an open connection, which may be a very difficult task.

However, in recent years the mobile industry has increased its support of a new `small cells` approach. These small cells are generic, mostly Linux based, home router-like devices developed by third party vendors. They are used by mobile carriers to provide connectivity in low coverage, remote areas, and use common backhaul connection (usually public internet lines) to connect directly to the core network. Attacking such a small cell may provide the attacker with a much more convenient way of gaining access to the backhaul and possibly into the core network. In the past few years, several cases were reported where such devices have been hacked or compromised. [4] [5] [6]

In fact, just recently, a Verizon Femtocell was reported to be hacked by security researchers that have demonstrated their ability to eavesdrop on cellular calls and text messages using the hacked device. [12]

Core Network

With the development of mobile network generations and the standardized architectures and protocols, there is a shift of the mobile core network from telephony infrastructure (such as Signaling System No. 7/SS7) to IP based carrier data technologies. For instance, the latest mobile network generations were designed with an IP centric approach and have dramatically increased the usage of standard IP based protocols like usage of Diameter protocol for signaling or SIP for the voice networks. Additionally there's the latest mobile generation – 4G LTE - which has designed the core network to be entirely IP based.

For an attacker, compromising the core network is considered the holy grail of mobile hacking, as it provides an attacker with virtually unlimited options and attack vectors.

Similarly to other computer networks, the internal network is much more difficult to defend than the perimeter. An attacker could realistically be faced with far less difficulties when executing malicious actions inside the core network than he would be facing on any of its perimeters.

Nevertheless, gaining access to the core network is the most difficult of all of the potential entry points, since these networks are closed and the perimeters are well-defended.

Still, the mobile core network has its security weak spots. For example, most mobile networks contain a unified management system (also referred to as OSS) that is used to control and provision network elements in the entire network (not only the core network). These systems require operators to open direct connections from the external networks. As with any application, these management systems might contain security vulnerabilities that will allow an attacker who is able to exploit them to escalate his access directly from the external network and pivot his attacks to the rest of the network.

External Networks

External networks currently pose the most realistic and accessible entry points to mobile networks. An example of an external network is the Internet, which is obviously very accessible and dynamic. Another example are semi-public networks, providing inner connections to other mobile carriers, in which there is little or no knowledge of the external network architecture or security design.

In the case of public networks, it is safe to assume that since they expose mobile providers to the greatest security threat, they are heavily protected. Common defenses that can be expected to be seen on such public network perimeters include firewalls, content inspection, load balancing and DPI (deep packet inspection) capabilities.

In the case of inner-provider connections, however, defense mechanisms may be far less restrictive, since external providers are considered trusted business partners. Connections are designed for sharing sensitive signaling and data information between core networks. Since data is business critical, mobile operators will not risk having it go through a tight security layer, which may potentially drop or alter the data.

An attacker may take advantage of this condition and use inner-provider connections to gain access from a small unprotected network area.

Availability Threats to Mobile Networks

Although mobile networks have a different design than traditional enterprise networks, they do share many common characteristics. Therefore, traditional denial-of-service attacks designed to misuse the lower protocol layers, such as IP, TCP and UDP are also relevant to mobile networks using these protocols. These attacks include, among others, SYN floods, Out-Of-State floods, UDP Floods, Abnormal packets floods, etc.

However, when it comes to the application layers, mobile networks introduce several unique attack vectors which are covered in the following paragraphs. The attacks are classified according to the following criteria:

- **Implementation Difficulty** – how complicated is it to execute an attack
- **Required Network Access** – the network point of access that is required for the attacker in order to effectively implement this attack
- **Potential Impact** – the potential impact to the mobile network if the attack is effectively implemented.
- **Overall Risk** – the estimated overall risk of the attack (taking into account its potential impact and difficulty of implementation)

Attach Floods

In order for a mobile device to communicate through a mobile network, a certain procedure must first take place. This procedure [7] includes a client connection request followed by an authentication procedure. If the client successfully completes this procedure, he is authenticated to the network, a connection channel is given to the device, and the client may then send and receive data.

In 2G, 3G and 4G networks, all data communication preformed after this initial authentication procedure is encrypted and digitally signed when sent over the radio network. This prevents eavesdropping, modifying or replaying of the data by third parties. However, the data flow preceding the authentication procedure (i.e. all data sent\received in stages 1-4 in the diagram below) contains no encryption or digital signing, since the trust required for implementing these protections is gained only after the authentication procedure (stage 5) is complete.

Difficulty of Implementation:
Medium-High

Required Network Access:
Modified Mobile Device

Potential Impact:
Medium-Full Network Outage

Overall Risk:
Medium

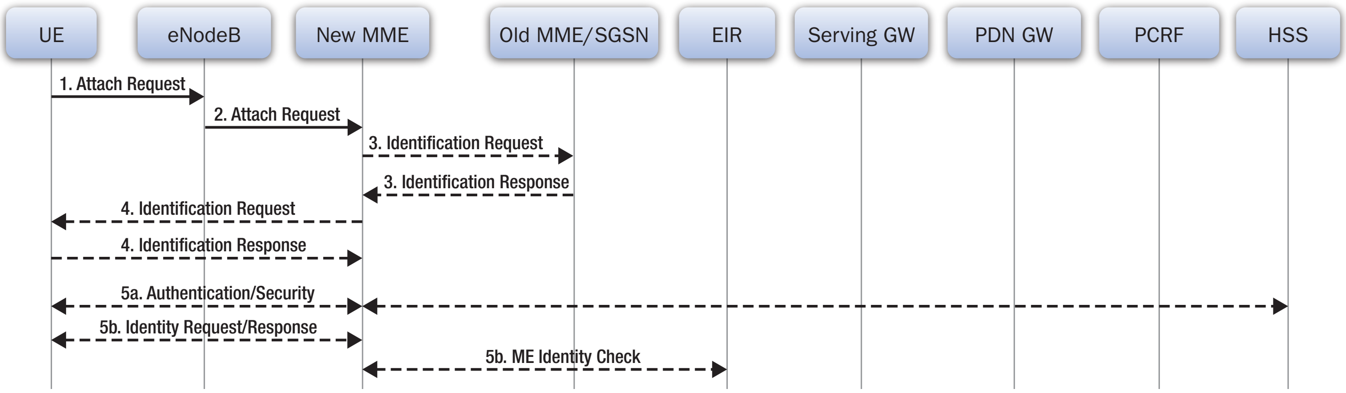
UE
eNodeB
New MME
Old MME/SGSN
EIR
Serving GW
PDN GW
PCRF
HSS
1. Attach Request
2. Attach Request
3. Identification Request
3. Identification Response
4. Identification Request
4. Identification Response
5a. Authentication/Security
5b. Identity Request/Response
5b. ME Identity Check

Figure 3 - LTE Mobile Device Attach Procedure (source: www.lteandbeyond.com)

This behavior could be misused by an attacker in order to create an attach flood. The attacker will send a high rate of fake attach messages to the network, causing high resource utilization on the backend signaling gateways (MME, HSS for 4G or MSC, BSC, VLR on 3G). This may cause anything from a local station outage to a local area network outage.

This attack can be further leveraged by also sending also fake IMSI`s to the network (which are the unique identifiers for the mobile subscriber stored in the SIM card). This would result in high resource utilization on the subscriber database (HSS for 4G or HLR for 3G) and may potentially cause a complete network outage.

It is important to understand that this type of an attack could be used to target roaming partners and could potentially cause outages on the remote carrier networks. This could be done by flooding using specific IMSI`s that belong to a specific remote provider. This will force the local core network to connect to the remote subscriber database for validation upon each request.

Another variation of this attack would be simply to execute since it does not require modifying mobile devices. An application could be loaded into a large number of mobile devices that will be used to consistently disconnect\reconnect to the network. This could be done, for example, by switching the ‘airplane’ mode on and off on smartphones. Such an approach would have a similar affect with a large number of devices connecting simultaneously to the network over and over again.

An application of this kind could be spread as a part of a mobile malware, or intentionally installed on mobile devices and coordinated from a “web hive” as part of a Hacktivist attack campaign.

It is important to note that with this attack variation it is easier for the mobile provider to detect and possibly block the offending devices and subscribers.

Paging Floods

The architecture of modern mobile networks allows mobile devices to have several active and idle states, which enable optimizing resource consumption on network radio links while still supporting the “always on” concept. From a network standpoint of view, idle devices do not require an open radio channel, while users do not notice any disconnection from the network and feel that they are always connected.

When a device is in an idle state and wants to re-initiate communication with the network - for example while placing a phone call - the procedure is similar to an initial network connection. The device sends a new network `Attach` signal and resumes its previous connectivity. However, when the network re-initiates communication with an idle device (such as in the case of an incoming call), a paging procedure is invoked. The network sends a “wake-up” signal to the device and the device then reinitiates its connection.

Difficulty of Implementation:
Very Easy

Required Network Access:
Public Internet Access

Potential Impact:
Medium-Full Network Outage

Overall Risk:
High

Although the mobile network keeps track of the device location, when it comes to handling an idle device, tracking is limited to the devices last reported location. Paging operations, therefore, usually begin with limited signaling on the last known base station, and gradually expand the paging signal area if no response is received from the device. Paging operations are not limited to phone calls, and the exact same procedure is followed when an idle device receives incoming data packets from an IP network such as the internet.

Attackers can exploit the paging behavior by flooding a public mobile device`s IP address with a large amount of traffic. This can be executed by launching a high capacity SYN flood, or a full TCP connection flood on a public network range, either maliciously or unintentionally due to a network scan operation. Assuming that at least 5-10% of mobile devices associated with these IP addresses will not be available at their previously reported location, the network will begin broadcasting paging signals to locate "missing" devices. These signals could dramatically stress the network, causing service disruptions and possibly even a network outage.

The main limitation of this attack vector is that in today's mobile networks most mobile devices are placed behind a NAT, so that their IP address is not reachable from the external network. However, there are several mobile network implementations that do not use NAT, and therefore might be vulnerable to such an attack vector.

Another mobile trend that may dramatically expose networks, and increase the surface for this attack is Machine to Machine (M2M) communication. M2M technology allows generic devices such as ATM`s, temperature sensors or tracking devices, in order to communicate with each other and with a main control center using a publicly exposed IP address, which may be used by attackers to overcome the NAT protections.

M2M is widely used. According to a market research paper [8] the number of cellular M2M subscriptions nearly doubled between 2010 and 2012- reaching 143.7 million M2M subscribers.

The following diagram illustrates an attack scenario:

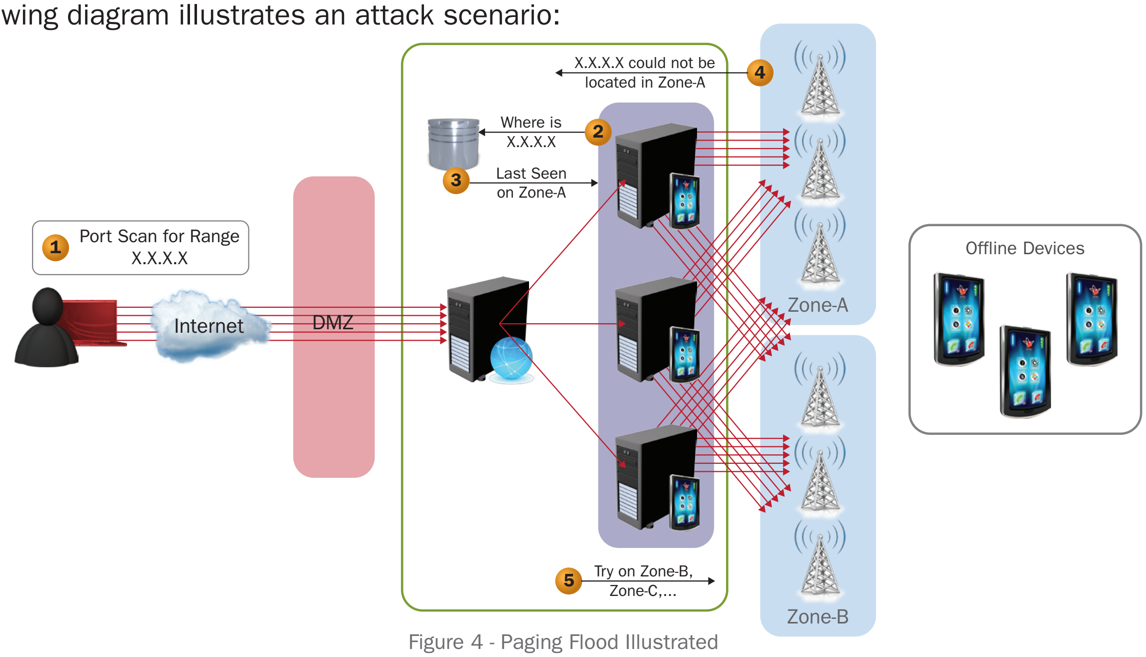
X.X.X.X could not be located in Zone-A
4
Where is X.X.X.X
2
3
Last Seen on Zone-A
1
Port Scan for Range X.X.X.X
Internet
DMZ
Zone-A
Offline Devices
5
Try on Zone-B, Zone-C,...
Zone-B

Figure 4 - Paging Flood Illustrated

1. An attacker locates a range of publicly accessible mobile device IP addresses and generates a massive port scanning on that range.
2. The network traffic generated by the attacker reaches the mobile network traffic anchors, which in turn query the mobile DB for the location of the target device.
3. The mobile DB responds with the last seen location of the mobile device.
4. The mobile network traffic anchors try to locate the device at the given location.
5. When no response is returned, the mobile anchors gradually increase the paging zone, thereby causing a massive paging operation that partially or entirely overwhelms the mobile network.

DNS Floods

As mobile network generations develop, DNS servers are gradually taking a more dominant role in overall network operations. In the latest 4G LTE architecture, the entire core network operations for both control and user planes is dependent on DNS services. In a case where DNS services are disrupted or stopped, the impact could be a full network outage.

Difficulty of Implementation:
Easy

Required Network Access:
Internet/Partner
Network Access

Potential Impact:
Medium-Full Network Outage

Overall Risk:
High

A common LTE architecture contains 3 DNS server roles:

1. **External DNS** – used by the mobile devices to resolve IP addresses in the external packet data networks (such as the internet).

   These DNS servers are usually located in a restricted network zone, such as a DMZ so that incoming connections from them to the core network is limited.

2. **Internal DNS** – used by the core network for resolving internal network element address. In LTE all Diameter based signaling and load balancing are dependent on these DNS servers.

3. **Intra-Operator DNS** – used by the core network to publish its network element addresses to other mobile providers, and by other mobile provider's to resolve local network elements addresses.

   This DNS server is crucial for supporting mobile device roaming abilities and inter-operator communications.

While the dependency on DNS servers is growing, DNS floods have become popular in the Denial-of-Service (DoS) world in the past few years. DNS attacks with rates reaching up to 50, 100 and even 200 Gig are quite common. If such attacks were directed at a modern mobile network, they would very likely cause a heavy stress and even force network services to be brought down for a significant period of time.

The latest "trend" in DNS based flood is called a reflective, amplified attack. This attack exploits the stateless nature of the DNS protocol and its lack of authenticity. In a reflective attack, attackers send specially crafted queries to DNS servers in such a way that the queries' source address is forged and set to contain the attacked address instead of their own. Since the servers response to the query may be up to 10 times larger than the requests, this may cause very high incoming traffic rates targeting the attacked server. This may cause it to freeze and in extreme conditions saturate the entire available network bandwidth.

In the internet domain, servers that can be used to perform such an attack are called `Open Resolvers` since they answer queries from unknown sources. In case the target of the attack is a local mobile network, DNS servers cannot deny requests; just like 'Open Revolvers' act in the internet domain, they will allow the attacker to use the mobile infrastructure to amplify the attack.

SIP Floods

IMS, which stands for the IP Multimedia Subsystem, is a network designed to replace the traditional voice networks used today. Its IP approach replaces all voice signaling protocols with SIP signaling and data carrying protocols with RTP and several other common IP based protocols.

Combined with richer abilities and functionalities in voice networks, IMS also exposes the voice network to new risks. SIP security is a well-researched field and most of the critical threats have already been mapped and studied. However implementing protections for these threats might prove to be a challenge for mobile carriers who wish to migrate their current voice networks to IMS due to the huge scales and decentralized nature of these networks.

Difficulty of Implementation:
Easy-Medium

Required Network Access:
Modified Mobile Device

Potential Impact:
Voice services outage

Overall Risk:
Medium-High

An example of the most typical and easily executed SIP flooding techniques is the SIP INVITE flood. An attacker can generate a high rate of SIP INVITE request, forcing the SIP Proxy or SIP server to create and store a session for each of the requests until server resources are depleted and a denial of service state is reached. There are many more attack vectors that can be executed on SIP based networks, with publicly-available automated attack tools that facilitate such attacks.

Signaling Storms

The amount of signaling used today by 4G LTE core networks is 10 times larger than it used to be on 3G networks. Since network operators concentrate most of their resources on upgrading user plane capacity in order to support user demand, control plane capacities are often neglected. This is because their traffic is not expected to grow at the same proportion as the user plane capacity.

Mobile networks that have not adapted to this rapid signaling increase may be vulnerable to signaling storms. Signaling storms cause a huge amount of signaling communication to simultaneously occur on the core network and can potentially cause a full network outage.

Difficulty of Implementation:
Easy-Medium

Required Network Access:
Mobile Device

Potential Impact:
Medium-Full Network Outage

Overall Risk:
High

Signaling storms can either be caused by malicious activity or by legitimate network usage. For example, chatty applications are known to be high signaling generators, since application developers' main concern is application functionality or battery savings, and little attention is placed on potential network stress.

Several cases of major operator network outages have been recorded and caused by signaling storms of Diameter traffic that took place due to heavy application usage or as a side effect of network maintenance operations [9] [10].

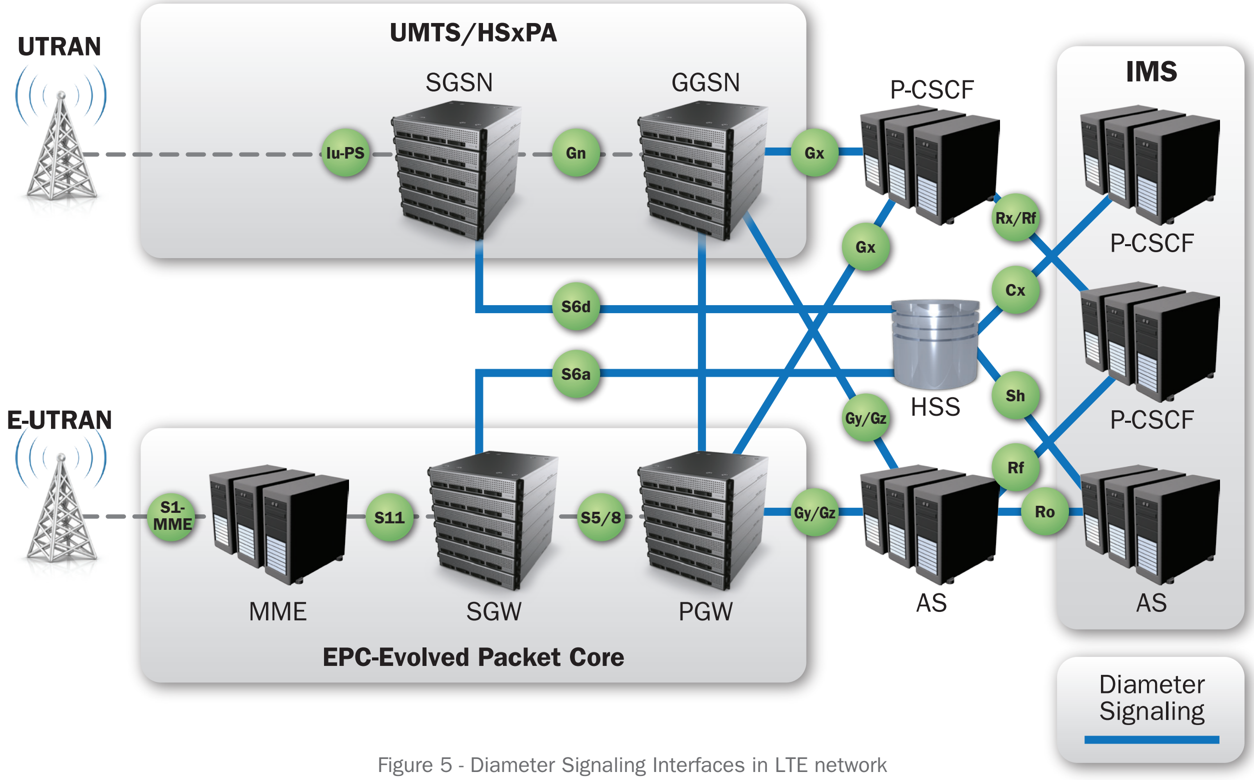

Figure 5 - Diameter Signaling Interfaces in LTE network

Summary

The following table summarized the attack vectors discussed in this report along with mapping of mitigation technologies that can effectively thwart them.

Attack Vector	"Traditional" DoS Attacks	Attach Floods	Paging Floods	DNS Floods	SIP Floods	Signaling Storms
DPI	·	·		·		
Behavioral Based Mitigation	·	·	·	·	·	·
VoIP Protections					·	
DNS Protections				·		
Signaling Proxies		·	·			·

Conclusion

There is no denying the growth, damage and popularity of Denial-of-Service (DoS) attacks in the enterprise environment. This has resulted in a significant investment in precautions and protections that were put in place to maintain service availability and mitigate attacks.

Mobile networks, which for so long have been happy and comfortable with their “walled garden” and sense of security it brings, must catch up with the network security world and face the growing threats. The shift towards IP-centric networks, combined with the rapid increase of knowledge and technology in the mobile world, are making mobile network exploitation easier, cheaper, and more accessible. Add that to the huge impact of bringing down a mobile network, and it is clear that a large increase of mobile network attacks can be expected in the near feature.

As the endpoints for mobile networks become open source devices, vulnerabilities will be exposed at a faster rate. Mitigation needs to be built into the environment and budgets need to be allocated. Above all, security needs the priority it deserves. To many organizations, communication is necessary, but to a Telco, it is beyond the point of critical.

Mobile networks should start planning a defense strategy that ensures network availability and providing protection from attacks originating both internally and externally.

We cannot predict the future – but we can prepare for it.

Works Cited

[1] grugq@coseinc.com, Base Jumping Attacking the GSM baseband and Base Station, BlackHat-USA-2010, 2010.

[2] S. Munaut, Targeted DOS Attack and various fun with GSM Um, DeepSec-2010, 2010.

[3] Harald Welte, hmw-consulting, Running Your Own Gsm Stack On A Phone, 27C3, 2010.

[4] M. Szczys, Poking at the femtocell hardware in an AT&T Microcell, 2012.

[5] Kevin Redon SECT, Technical University of Berlin, Hacking Femtocells a femtostep to the holy grail, 2010.

[6] T. H. Choice, The Vodafone Access Gateway / UMTS Femto cell / Vodafone Sure Signal, thc.org, 2011.

[7] B. Barton, LTE attach procedure, 2012.

[8] Jan ten Sythoff, Pyramid Research, Cellular M2M Connections - An Analysis of Growth Drivers, Market Segments and Operator Approaches, Pyramid Research, 2013.

[9] Caroline Gabriel rethink-wireless, DoCoMo demands Google’s help with signalling storm, 2012.

[10] Kevin Fitchard, Gigaom, Why are mobile networks dropping like flies?, 2012.

[11] Michelle Donegan, Light Reading, Docomo counts cost of signaling storm, 2012.

[12] J. Finkle, Researchers hack Verizon device, turn it into mobile spy station, New York: Reuters, 2013.