HEAVY READING

White Paper

Mobile Security: Responsibilities & Opportunities for Operators

Prepared by

Patrick Donegan
Chief Analyst, Heavy Reading
www.heavyreading.com

on behalf of

CISCO™

www.cisco.com

December 2015

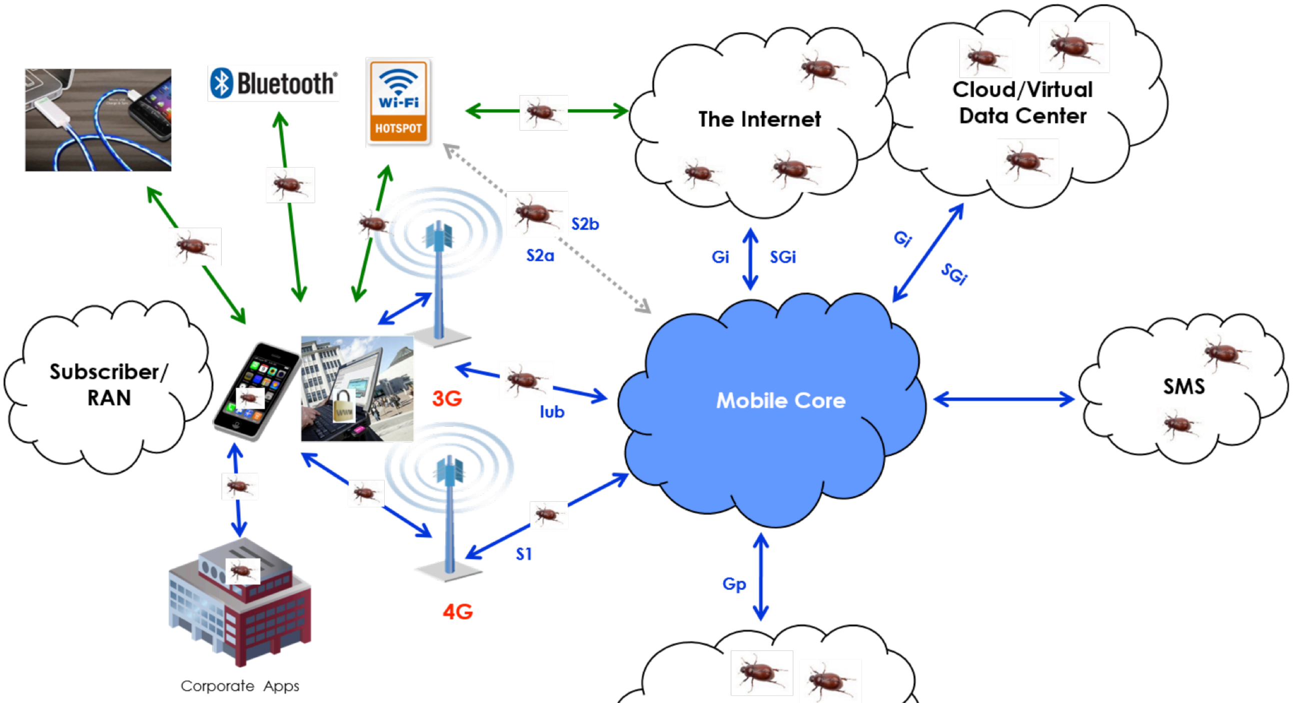
An Operator's View of Mobile Security Threats

Throughout the 2G and early 3G eras, the security challenges faced by mobile operators were on a far smaller scale than they are today. The built-in 3GPP security between the handset and the BSC or RNC provided an excellent solution to the risk from eavesdropping. Yes, there was some fraud – relating to SIM cards and roaming, for example – but that fraud tended to impact the operator by using network resources without paying for them. Typically, the operators' customers were wholly unaffected, the operator's ability to continue billing customers was unaffected, and the operator's reputation and brand were likewise unaffected.

The rollout of 4G is rapidly accelerating the transformation in mobile networks, application and services to an ecosystem dominated by IP. Inevitably accompanying that transition is a huge spike in security threats and vulnerability to those threats. As 4G connectivity, capacity and data consumption scales up, mobile operators now no longer dare to assume the same relatively static, relatively small-scale threat landscape that they once could.

As shown in **Figure 1**, there are a variety of different attack vectors that are routinely used to launch attacks at mobile operators, their business partners in the information and communications technology (ICT) value chain, and their end-user customers.

Figure 1: Mobile Operators Faces Security Threats From Multiple Sources



Source: Heavy Reading

As shown, malware and malicious traffic can come in via any of the many wired and wireless access networks that connect smartphones and other devices to the network nowadays; via the Internet or cloud services over the Gi and SGi interfaces; via SMS; and via roaming partners. According to Heavy Reading's October 2014 Mobile Operator Survey on Mobile Security, an average of 4 percent of a mobile operator's daily traffic is now malicious traffic. In the case of WiFi access networks, attacks target voice over WiFi (VoWiFi) applications as well as data applications.

An Understanding of Security Threats, Solutions & Use Cases

Today's mobile operators, most of which are looking to aggressively scale up their 4G networks, need to devise and implement their security strategies with three primary requirements in mind. As elaborated throughout this paper, these are:

- A granular understanding of, and visibility into, the variety of security threats that face them and their customers.
- An appreciation of the security practices and technologies needed to protect their infrastructure and their customers, their revenues and reputation.

- An understanding of some of the key mobile security use cases for protecting current revenue streams – and growing new ones.

At a high level, there are two types of cyberattack that are most threatening to a mobile operator's business performance. These are attacks that cause network outages or degradations in service; and attacks that steal or expose customer information, either to steal money – or other valuable data – or to cause reputational damage. A mobile operator's failure to defend its infrastructure and its customers against these primary security threats significantly increases its risk of revenue loss and damage to brand equity.

The revenue loss can be direct, when the attack causes the operator's ability to bill for service to be suspended; or it can be indirect, when the attack doesn't affect the ability to bill for service, but does impact the user experience badly enough to damage the operator's own reputation and brand as a trusted service provider, and hence drive higher customer churn rates.

Fraud & Theft or Exposure of Customer Data

Heavy Reading research has consistently shown that the security attacks that mobile operators worry about most are those that steal or expose customer data and those that impact network performance. In our October 2014 survey, exposure or theft of user data scored 4.1 on a scale of 1 to 5, where 5 was the highest level of concern.

There are three main ways in which attackers – whether they are external attackers or rogue insiders – can use the mobile network to carry out fraud as well as steal or expose sensitive customer information. They can directly steal money (e.g., via a fraudulent text triggering a response to a premium rate service); they can steal information (with a view to fraudulently stealing money, stealing value such as intellectual property, or making money from that information without the user's authorization); and they can steal and expose information solely with the intent of causing harm to the victim.

These types of attack tend to take one of three different forms:

- Malware that exploits or exfiltrates from data repositories in the operator's network, in the cloud, on a user's smartphone, or in the enterprise network (leveraging unauthorized access to the user's smartphone).
- Malware that resides on a device or network element and can eavesdrop on user communications by recording, storing and forwarding real-time communications sessions (otherwise known as spyware).
- Breaking into the flows of user plane traffic on the interfaces between mobile network domains in order to intercept and eavesdrop on real-time voice and data communications.

2015 has seen a number of high-profile incidents of attackers gaining access to the bank account details of the customers of a number of service providers, ranging from relatively small ISPs to large Tier 1s.

Why the Operator Needs to Step Up & Lead

Resistance to spending more than the bare minimum on protecting against fraud and exposure of customer information is still remarkably common among some mobile operators today.

This hesitation is sometimes rooted in the financial constraints that many operators are faced with. Sometimes it is also rooted in an important but unflattering truth about some consumers and some enterprise IT managers – namely that if the end user doesn't contribute to their own security, then much of the operator's efforts on their behalf are cost-prohibitive, if not wasted.

After all, the mobile operator can't control a consumer's propensity to visit high-risk Websites. The operator can't control individual smartphone security settings or security patch updates, either. Nor can it directly control the enterprise's own internal security processes with respect to the bring-your-own-device (BYOD) policy or access to corporate databases from 3G- or 4G-enabled devices (unless it is formally contracted to do so).

But clinging to this legacy mentality that is only applicable to one segment of customers, and putting that at the heart of a strategic reasoning for a minimalist network security policy is increasingly open to question when exposed to scrutiny. And there are three critical reasons for this:

- **The operator is impacted, irrespective of who is to blame.** Whether or not the individual customer is directly responsible for succumbing to an attack, the operator typically gets the blame and suffers the additional calls into call centers, negative social media coverage and churn impact anyway. Some operators have seen as much as 30 percent of the value of their companies wiped out following high-profile attacks on customer data. The question isn't "Who is to blame?" for attacks, but rather, "What are the short- and long-term consequences? Do they matter? And how can they be prevented?"
- **The burden of cyberattacks aren't equally shared across users and operators.** For example, if four operators in a given market are being targeted with messaging spam, and one takes effective action to block it, attackers will just direct more of their attacks at the operators with the weaker defenses.
- **Segments of customers are willing to pay for security services.** That might be a dollar or two per month for additional malware detection and mitigation service, or as a unique service available only as part of a premium subscription. Enterprises will often pay for mobile security as a managed service. Moreover, many major new mobile broadband revenue opportunities derive revenue not from end users but from third parties, as in the case of many m-health and m-commerce use cases. Some industry verticals are potentially willing to spend hundreds of millions of dollars on new mobile-enabled business models. But where they perceive a risk to their own or their customers' information, they will only partner with those mobile operators that offer the most advanced network security. They will decline to partner with those that only offer the bare minimum.

Attacks That Cause Outages & Degradations

In our October 2014 survey, mobile operator respondents scored network outages or service degradations their highest concern, ranking an average of 4.3 on a scale of 1 to 5. Where outages are caused by adversaries, they are usually the consequence of distributed denial-of-service (DDoS) attacks. These can be volumetric, whereby large volumes of requests are sent directly at an end target domain, overwhelming it; or they can be leveled at the application layer, whereby relatively low volumes of malicious traffic overwhelm application servers' supporting services like SIP or DNS.

DDoS Impacts on a Mobile Network? Really?

As of today, the risk of DDoS attacks triggering a mobile network outage is lower than the risk posed by equipment failure, a fiber cut, or a misconfiguration error. With relatively few public IPv4 addresses to draw upon, mobile operators also translate their customers' IP addresses via network address translation (NAT) gateways. The fact that attackers can't see mobile users' IP addresses serves as a useful security buffer (although that too may change once mobile operators roll out IPv6).

None of this means that the risk posed to mobile operators by DDoS attacks is negligible. On the contrary, Heavy Reading research shows that the risk is already significant and growing. In our October 2014 survey, 36 percent of mobile operator respondents said that their company had suffered a network outage lasting at least an hour twice or more during the previous 12-month period. Only 40 percent stated their company had avoided any such customer-impacting incidents.

As shown in **Figure 2**, almost two thirds of the network security experts in our October 2014 survey of mobile operators cited the DNS infrastructure as having been impacted by DDoS attack traffic during the previous 12 months – more than any other network domain.

Figure 2: Mobile Network Domains Most Impacted by DDoS Attacks

Domain	% of Mobile Operator Security Experts Citing a DDoS Performance Impact
DNS	63%
Services (web, email, SMSC)	47%
Subscriber devices	42%
NAT gateways/firewalls	32%
HLR, HSS, AAA system interfaces	32%
GGSN/SGSN	32%
Packet core routers	16%
RAN	16%

Source: Heavy Reading's 4th annual Mobile Network Security Survey (October 2014); respondents qualifying as security experts #19

Almost half of respondents said their Web, email, SMSC and MMSC infrastructure had been affected, and over 40 percent said their customers' devices (typically smart-phones) had been affected. While attacks on the HLR/HSS and AAA systems are less common, the impact of a successful attack on these elements can be much bigger.

Until now, the bulk of the disruption to mobile service caused by malicious traffic has not come about as a result of attackers expressly setting out to take down the mobile operator's own infrastructure. Rather, it's come from the operator serving as an unwitting conduit, allowing malicious traffic to transit its network and cause disruption.

From an attacker's perspective, there is, in aggregate, even more incentive to steal data from consumers and businesses than to disable the operator itself. Service providers that unwittingly enable malicious traffic to reach their customers' end points imperil their business as much as do attacks on their own infrastructure.

Smartphone Botnets Are a Real Threat Now

Increasingly, a variety of different attack types rely on botnets to distribute malware. Up until recently, the threat of mobile devices being infected by botnet malware has been confined to 3G- or 4G-connected laptop users. This is because the traditional focus of attackers was almost exclusively on the Windows OS, so only 3G- or 4G-connected laptops were vulnerable to becoming part of a mobile botnet.

The last 18 months have seen a steep change in the pervasiveness and quality of smartphone botnet software. The November 2014 discovery of "NotCompatible.C" for Android was a landmark in smartphone botnet software development because it supported sophisticated command and control and encryption to avoid detection – attributes that had traditionally only been seen in Windows-based botnet malware.

More recently, in September 2015, some 650 million Chinese smartphones were found to have downloaded malicious software from an online advertisement. The malicious software then managed to generate a staggering 4.5 billion requests targeted at a website, in an effort by attackers to take it down. The threat is also mounting with new LTE and LTE-A roadmaps, which are enabling mobile devices to access faster and faster speeds, including on the uplink, which is key for launching DDoS attacks.

There are at least three reasons why mobile operators need to prevent their subscribers from becoming part of a botnet to distribute DDoS and other types of attacks:

- **Botnet software residing on a smartphone can cause a deterioration in that smartphone's performance.** This can result in inferior application performance for the infected user. It can also reduce battery life, increasing the risk of the end user becoming unable to access revenue-generating services.
- **Malicious traffic originating from an infected subscriber can cause congestion in the mobile operator's own network.** This can impact the user experience of other customers and require network capex to be brought forward.
- **The spread of botnet software across the subscriber base increases the mobile operator's vulnerability to IP address blacklisting.** As already mentioned, the unique pooling of public IP addresses in the mobile network does have security advantages. In the mobile network, IP addresses are dynamically shared by multiple users, allowing connections to be set up and then quickly torn back down again. Hence, if one of a mobile operator's public IP addresses gets blacklisted for being part of a botnet, it isn't just one user that has their Internet access blocked; it may be many users simultaneously.

Multi-Layered Security Practices & Technology

Meeting mobile security challenges in the 3G and 4G eras requires a multi-layered approach – securing multiple devices and multiple interfaces against multiple threats in the operator's own network, in the cloud, and on the end customer's devices.

An optimal strategy must take account of the increasing sophistication, variety and volume of threats. It also must be done within the operator's financial constraints. These typically leave little scope for increasing headcount in the operator's security team, especially given the global shortage of network security experts that is enabling those with the right skills to command high salaries. And it needs to be done in a way that allows the security team to focus more of its time on dealing with high-risk threats. It therefore must be done in a way that introduces as much real-time or near-real-time automation as possible into the operator's incident management processes.

The following are among the key capabilities that mobile operators need to have available to them to align with these emerging security requirements.

Threat Intelligence, Network Visibility & Anomaly Detection

One of the key things that separates leaders from laggards in mobile security is the ability to detect malicious packets as well as behavioral indicators of compromise. It's remarkable how many CTOs and CEOs in these companies still have very limited visibility into exactly how much – let alone exactly what type – of malicious traffic they are actually carrying. Access to the best threat intelligence matters – in particular from where the intelligence is drawn, how up-to-date it is, and how it is made accessible.

The majority of well-known security threats can still be identified by recognizable signatures, so identifying these signatures continues to be very important. But a growing minority of the most effective attacks don't have a recognized signature. They are customized by criminal hackers to exploit weaknesses in specific infrastructures and are designed to avoid detection. An anomaly detection capability that understands an operator's baseline traffic profile, and then accurately flags high-risk deviations from that baseline, is becoming a must-have for security-savvy mobile operators.

Threat Mitigation

Multi-layered mitigation requires defensive measures being put in place before, during and after an attack. A lot of preventive mitigation has to do with applying the right network security policies and the right internal security procedures in terms of employee and partner access to network resources.

Real-time mitigation can mean anything from removing malware from infected smartphones to deciding to direct an entire traffic stream to the operator's own scrubbing center or out to a scrubbing center in the cloud. The challenge for operators is to procure solutions that can automate an increasing proportion of mitigation responses in real time – without generating false positives or negatives that flag legitimate traffic as illegitimate, or vice versa. After an attack, the mobile operator needs the capability to quickly and accurately trace the root cause.

Traffic Encryption

For 2G and 3G networks, most of what the mobile operator needs in terms of encryption is built into standardized 3GPP equipment. This provides encryption from the handset across the RAN to the BSC or RNC, which is pretty secure, deep in the network. Today's environment is very different: In 4G, 3GPP encryption terminates at the eNodeB. That leaves clear text from the eNodeB across the backhaul network unless the operator implements its own encryption. Network encryption has moved on in a lot of other ways too.

End-user requirements for encryption create a growing market opportunity for operators to sell encryption as a service as one strand in a strategy for monetizing security. At the same time, adversaries have figured out that encrypting their attack traffic is a smart way to avoid detection. Hence, encryption solutions that operators can provide customers that also allow the operator to inspect, decrypt and re-encrypt with the customer's agreement are becoming increasingly valuable.

Open Networking Principles

The concept of multi-layered security is complimentary to the evolution of mobile networks, both in terms of the all-IP Evolved Packet Core (EPC) as well as the emerging transition toward highly orchestrated software-defined networks (SDN). These changes in the networks are evolving in parallel with a transition from a focus on loosely integrated, siloed security services, using multiple dedicated hardware platforms from many security vendors, toward more tightly integrated security services that protect data consistently across physical, virtualized and cloud architectures.

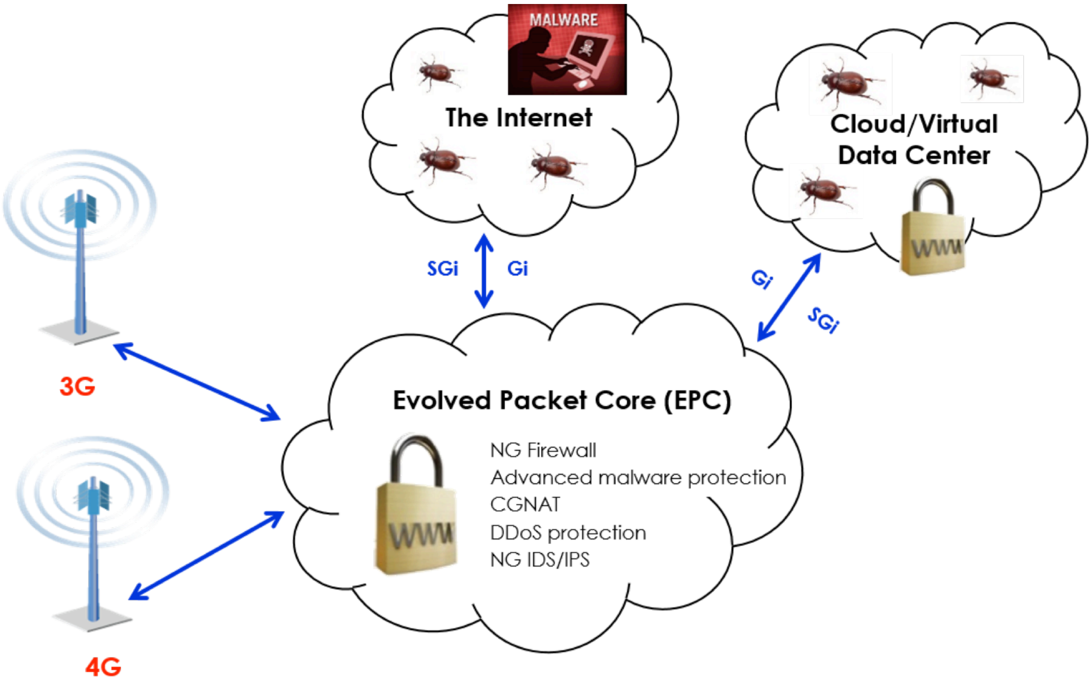
Going forward, game-changing trends in telecom networking such as SDN and network functions virtualization (NFV) have the potential to enable a more flexible, lower-cost, and potentially more robust multi-layer approach to mobile security. Virtualization can be transformational for 4G networks, but it will be foundational for 5G.

Key Mobile Security Use Cases for Operators

The breadth of visibility and insight that it can potentially have, combined with the unlimited processing capacity it can draw upon without impacting the user experience, makes the operator's own network infrastructure the workhorse for any mobile security strategy.

The single most important mobile security use case consists of protecting the mobile packet core or the EPC in 4G. As shown in **Figure 3**, this is the heart of the mobile network where it meets the open Internet via the Gi (3G) and SGi (4G) interfaces. These interfaces are where the operator's infrastructure and its customers are most vulnerable to cyberattacks.

Figure 3: Securing the Gi & SGi Interfaces



Source: Heavy Reading

Use Case #1: Securing the Gi & SGi interfaces

Because of this exposure to the open Internet, the Gi interface has always been a key focus for mobile security. Indeed the term "Gi firewall" was first coined at the time of the first GPRS and CDMA 2000 deployments 15 years ago. With 3G and now 4G, the original case for a Gi firewall is unchanged in one respect. Our October 2014 survey showed that 74 percent of respondents cited the Gi interface as the place where most DDoS attacks on the mobile network are targeted. No change there: The Gi is still the point of greatest vulnerability.

But in terms of what's needed to protect the mobile packet core today, the requirements have changed. Firewalling is required at the Internet peering point where the mobile operator interfaces to Internet transit providers and is vulnerable to threats contained in that traffic. It is also required on the Gi service LAN to allow traffic steering and protect against threats from the mobile operator's own subscribers.

Firewall access rules are still critical, but they are no longer sufficient in today's environment. Integrated threat defense against sophisticated and dynamic malware is also required, as is a fundamental shift in tactics to better automate incident response after an attack. In addition to next-gen firewall application visibility and control, mobile packet core or EPC protection also requires a variety of other tightly integrated security services, such as the latest in intrusion prevention, which offers advanced threat correlation and visibility, DDoS attack mitigation and automated advanced malware detection, mitigation, containment and remediation, as referenced above.

These security services must be efficiently orchestrated, with intelligence in the network to not re-inspect data where that is not required. The intelligent chaining of security services must occur both in the context of traditional environments as well as where SDN and VNF are employed.

Scalability requirements have also changed beyond recognition. The mobile data revolution of the last few years has generated a colossal increase in both total and malicious traffic. Coping with that does not simply require platforms that are carrier-grade, as compared with earlier enterprise-grade Gi firewalls. Security traffic inspection and network analysis software must have tremendous scalability to address future increases in mobile traffic. Traffic inspection must also scale while maintaining consistent policy across dedicated hardware, virtualized COTS platforms, and the bursting of workloads into service provider cloud networks. This capability also enhances new revenue opportunities with managed security service providers (MSSPs).

Use Case #2: S1 & X2 Security Using the 3GPP SEG

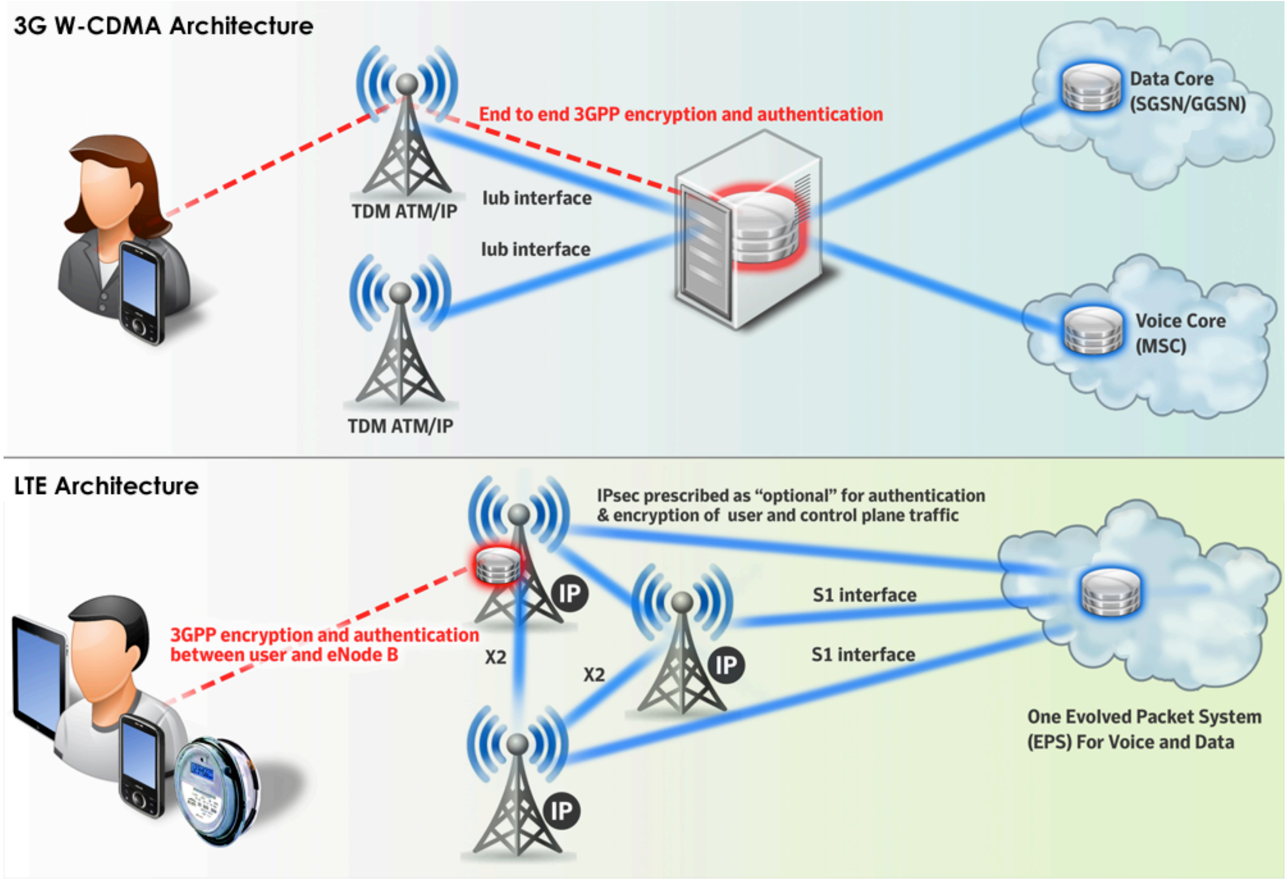
In the 4G network, the termination of 3GPP encryption at the eNodeB has already been noted, as has the fact that this leaves clear text from the eNodeB across the backhaul network, as shown in **Figure 4** (next page).

Operators need to think about protecting both the S1 and X2 interfaces, which are the interfaces between the eNodeB and the core, and between eNodeBs, respectively. First, the 3GPP's 4G Security Gateway (SEG) provides for IPsec encryption and PKI authentication to be used. Without it, the network is susceptible to eavesdropping on customer communications, malicious attacks on the operator's network infrastructure and malware targeting customer devices.

The risk increases with the rollout of 4G small cells, which compared to conventional macro and micro cells have much less physical security to prevent access and tampering. IPsec encryption prevents attackers from gaining access to the user and management traffic. PKI authentication ensures that only eNodeBs that have an approved vendor certificate embedded in them are authenticated onto the network. With smartphones increasingly vulnerable to infection by malware, including botnets, and with the Internet of Things set to rapidly drive up the number of connections, mobile operators also need to consider deploying some subset of the same threat protection they use to deploy the Gi to work in protecting the S1 and X2 as well.

A third use case for mobile security is securing the subscriber roaming interface. This leaves the mobile operator vulnerable to threats such as billing fraud from roaming partners that don't implement proper billing policies, unless the operator protects against it.

Figure 4: S1 & X2 Security



Source: Heavy Reading

Conclusion

Mobile operators need a clear understanding of the types of security threats they face, the approaches and techniques needed to combat them, and the use cases for deploying those defenses. Threats that target network uptime and performance and threats that target the exposure of customer data represent the biggest risk to the mobile operator in terms of their direct and indirect revenue impacts.

Mobile operators require an increasingly sophisticated suite of threat detection and mitigation capabilities that address anomalies as well as known signatures. These defenses need to provide a platform for increasing the level of automation in the incident management process. And they need to leverage the potential created by the transition to a more flexible, scalable, software-centric network driven by NFV and SDN. Protection of the Gi/SGi interface as well as the Gi/SGi service LAN, and protection at the S1/X2 interfaces, are two of the primary mobile security use cases as operators scale up their 3G and 4G networks.

About Cisco

Cisco (NASDAQ: CSCO) is the worldwide leader in IT that helps companies seize the opportunities of tomorrow by proving that amazing things can happen when you connect the previously unconnected. For ongoing news, please go to http://thenetwork.cisco.com.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. A listing of Cisco's trademarks can be found at www.cisco.com/go/trademarks. Third-party trademarks mentioned are the property of their respective owners.