

---

---

# Detecting SS7 Attacks in the Telecommunication Infrastructure Using SIEM

---

---

Master Thesis  
Marco Cappellari

Department of Electronic Systems  
Fredrik Bajers Vej 7B  
9220 Aalborg Ø



# AALBORG UNIVERSITY

## STUDENT REPORT

### Department of Electronic Systems

Fredrik Bajers Vej 7B

9220 Aalborg Ø

<http://www.aau.dk>

**Title:**

Detecting SS7 Attacks in the Telecommunication Infrastructure Using SIEM

**Theme:**

Masters Thesis

**Project Period:**

Spring Semester 2025

**Project Group:**

1049F

**Participant(s):**

Marco Cappellari

**Supervisor(s):**

Tatiana Kozlova Madsen

Niels Tobias Nørgaard Svendsen

**Page Numbers:** 45**Date of Completion:**

June 3, 2025

**Abstract:**

Every year multiple attacks are carried out over the global network infrastructure without being detected due to the trust based nature of the SS7 protocol suite. This thesis aims to develop an automated identification mechanism using a SIEM system to detect two location tracking attacks carried out using SS7 protocols. Network simulations of different operators interacting with each other have been used to generate network logs consisting of isolated attacks, but also a realistic scenario in which regular traffic was also being exchanged. Based on the attack patterns that have been identified, a search for each attack was first defined, then tested on the isolated scenario, and finally on the realistic simulation to determine their accuracy. The execution of such searches has then been configured to periodically run each day and notify the appropriate roles in case of a detection.

# Contents


<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Problem Analysis</b>	<b>3</b>
2.1	SIEM Systems . . . . .	3
2.1.1	Splunk . . . . .	5
2.2	Threats over the Telecommunication Infrastructure . . . . .	7
2.3	Infrastructure of a Telecommunication Network . . . . .	8
2.3.1	Elements of the GSM Infrastructure . . . . .	9
2.4	Signaling System No.7 . . . . .	11
2.4.1	Network Stack . . . . .	12
2.4.2	Attack Vectors . . . . .	13
2.4.3	Attacks over MAP . . . . .	15
<b>3</b>	<b>Detection of Attacks</b>	<b>18</b>
3.1	Issues Faced and Possible Solutions . . . . .	18
3.1.1	Compliance with Regulations . . . . .	18
3.1.2	Data Anonymization . . . . .	19
3.1.3	Simulators and Datasets . . . . .	20
3.2	Generation and Import of Network Logs . . . . .	21
3.2.1	Import of Data in Splunk . . . . .	22
3.3	Identification of Attack Indicators . . . . .	23
3.3.1	Any Time Interrogation . . . . .	24
3.3.2	Send Routing Info . . . . .	26
<b>4</b>	<b>Creation of the Correlation Searches</b>	<b>29</b>
4.1	Definition . . . . .	30
4.2	Mapping . . . . .	31
4.3	Timing . . . . .	32
4.4	Adaptive Response Actions . . . . .	33
4.4.1	Structuring the Email Notification . . . . .	34
4.4.2	Structuring the Notable Event . . . . .	35

Contents	iii
<b>5 Conclusion</b>	<b>38</b>
5.1 Input Sources and Flow of information . . . . .	40
5.2 Acknowledgments . . . . .	41
<b>Bibliography</b>	<b>42</b>
<b>A Searches for Displaying Purposes</b>	<b>a</b>

# Preface

Aalborg University June 3, 2025

The SIEM system developed by Splunk, called Splunk Enterprise Security, is referred to as "Splunk" throughout the report.

A handwritten signature in black ink, appearing to read 'Marco Cappellari', is positioned above a horizontal line.

---

Marco Cappellari  
<mcappe23@student.aau.dk>

# Abbreviations

A list of the abbreviations used in this report, sorted in alphabetical order:

2G	Second Generation
4G	Fourth Generation
5G	Fifth Generation
ATI	Any Time Interrogation
BS	Base Station
BSC	Base Station Controller
CDN	Content Delivery Network
CFCS	Center For Cyber Security
CIA	Confidentiality, Integrity and Availability
CSV	Comma-Separated Values
CSR	Cell Switched Routers
DDoS	Distributed Denial of Service
DoS	Denial of Service
EDR	Endpoint Detection and Response
EIR	Equipment Identity Register
ENISA	European Union Agency for Cybersecurity
GDPR	General Data Protection Regulation
GSM	Global System for Mobile Communications
GSMA	GSM Association
gsmSCF	GSM Service Control Function
GT	Global Title
HLR	Home Location Register
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity
IS	Information Systems
ISP	Internet Service Providers
ISUP	Integrated Services Digital Network User Part
IT	Information Technology
IoT	Internet of Things

JSON	JavaScript Object Notation
JSS7	Java SS7
MAP	Mobile Application Part
MSC	Mobile Switching Center
MSISDN	Mobile Station International Subscriber Directory Number
MTP-1	Message Transfer Part 1
MTP-2	Message Transfer Part 2
MTP-3	Message Transfer Part 3
NDR	Network Detection and Response
OSI	Open System Interconnection
PDML	Packet Description Markup Language
PLMN	Public Land Mobile Network
PSI	Provide Subscriber Info
RAN	Radio Access Network
SCCP	Signaling Connection Control Part
SCTP	Stream Control Transmission Protocol
SIEM	Security Information and Event Management
SIM	Subscriber Identity Module
SMS	Short Message Service
SMSC	Short Message Service Center
SoC	Security Operations Center
SOAR	Security Orchestration, Automation, and Response
SP	Signaling Point
SPL	Search Processing Language
SQL	Search Query Language
SRI	Send Routing Info
SS7	Signaling System No. 7
TCAP	Transaction Capabilities Application Part
UE	User Equipment
URL	Uniform Resource Locator
VLR	Visitor Location Register
XML	Extensible Markup Language

# Chapter 1

## Introduction

With the development and introduction of the Fourth Generation (4G), and the Fifth Generation (5G) of mobile cellular network infrastructure, the focus on the needs for mobile communications began prioritizing the data services, such as the access to the Internet, over traditional voice communication [1]. 5G and its future successor are, among others, focusing on scenarios achieving ultra reliable, low latency and high data rates communications, which would enable a greater number of devices to interact with each other [2].

The integration between the functionalities of a traditional mobile network, found in voice calls and Short Message Service (SMS), in relation to the access to the Internet, posed new challenges for the telecommunication industry. As the Information Systems (IS) governing the interaction processes in the cellular network are becoming more and more Internet dependent, new challenges in terms of reliability, efficiency, but also security arose [3]. If considered that the complete implementation of 5G will expose the entirety of the cellular core network of a telecommunication provider to the Internet, a network section typically isolated from it, such infrastructure will now become a possible target for malicious attacks [4].

The Internet exposure of ISs in the critical infrastructure is pushing the operators in the telecommunication sector to develop an Information Technology (IT) driven approach, which aims to apply IT security best practices to the protection of core networks [5]. This implies the adoption of models such as the Confidentiality, Integrity, and Availability (CIA) triad for securing them. The pillars composing the CIA triad are used as a starting point for the definition and implementation of protection mechanisms, aiming to improve the security posture of ISs in the networks [6]. Confidentiality aims to prevent unauthorized, i.e. not trusted, parties from accessing specific information in the communication between devices and the network infrastructure. Integrity focuses on the protection of the contents from changes, ensuring the information is valid and has not been altered. Availability aims to keep the systems operational at any time when requested by an authorized user.



These objectives are typically achieved through the Security Operations Center (SoC), a team responsible for performing incident response and management of security threats.

When referring to network infrastructure, the SoC typically makes use of three different systems to perform their work. The first is found in the Security Information and Event Management (SIEM), responsible for the analysis and aggregation of logs produced by the ISSs. The second is the Endpoint Detection and Response (EDR), which actively detects malicious activities in the network. The third is the Network Detection and Response (NDR), which dynamically analyzes the network activities in search of potential attacks [7].

Given that SIEM systems are traditionally used in an IT setting, and considering the changes undergoing in the telecommunication sector related to the implementation of the requirements for the new generations of cellular technologies, as well as their increasing association with IT practices, the following initial problem arises:

*How can the use of a SIEM system improve the handling of security vulnerabilities in the cellular network infrastructure of a telecommunication operator?*

## Chapter 2

# Problem Analysis

*This section provides a theoretical background of the elements and knowledge that have been used during the realization of the thesis and concludes by defining the final problem statement. It begins with section 2.1 by further defining the functionalities that characterize SIEM systems, focusing, in subsection 2.1.1, on analyzing the structure of Splunk, its search language, and making considerations related to its license models. Section 2.2 provides an overview of the attack trends for the telecommunication sector in recent years. A general description of the network infrastructure of a telecommunication network is then provided in section 2.3 with a focus on the Second Generation (2G) networks. Lastly, section 2.4 focuses on the Signaling System No. 7 (SS7) protocol suite by first describing its protocol stack, and later describing the attacks that can be performed on it, with a focus on the Mobile Application Part (MAP) protocol.*

### 2.1 SIEM Systems

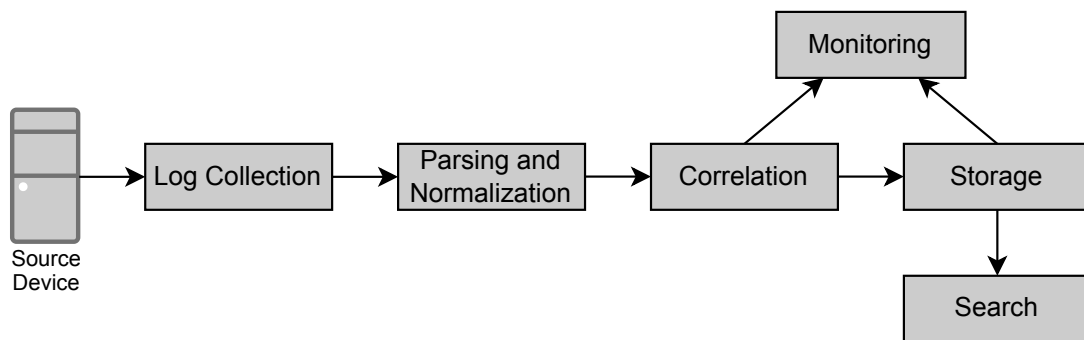
The logging of events in the enterprise networks is a security best practice standard, other than a requirement for compliance with cybersecurity frameworks or legislation [8]. Virtually all services running over a machine can be instructed to generate log records that keep track, in a timely manner, of events violating defined policies. As the information can be generated by multiple sources executing on devices operating from different network locations, SIEM systems have been developed as a centralized place in which events can be stored and analyzed [9].

SIEM systems are therefore used to aggregate in a common and immutable structured format the log flows generated by the entities in the network and to display, through specific views, the correlations between events, found through custom filters or rules, that may indicate the presence of potential threats [10]. This ability to highlight patterns of malicious behavior not only supports the work of the SoC but also aids roles like network operations, as the data related to the health of the services can also be logged to be later used to detect possible faults and their causes in the infrastructure. The services provided

by SIEM systems can therefore be classified in [10]:

- Log management, achieved by configuring the services in the infrastructure to forward the logged events to the SIEM system.
- Compliance, enabled via log management through the generation of periodic audits and reports. This allows the SoC to obtain an overall view of the status of the infrastructure, but also comply with cybersecurity frameworks and legislation.
- Correlation of events, realized through the contextualization of the events, i.e. combining the logs of different sources related to a particular system or entity before generating an alert.
- Active Response, enacted by the possibility of a SIEM system to automatically execute customized notification mechanisms in response to a detected event.
- Endpoint Security, enabled through an update on the way a service behaves based on the events analyzed by the SIEM system.

Figure 2.1 showcases the order of actions and operations performed by a SIEM system for the ingestion and elaboration of data:



**Figure 2.1:** Order of actions and operations performed by a SIEM system.

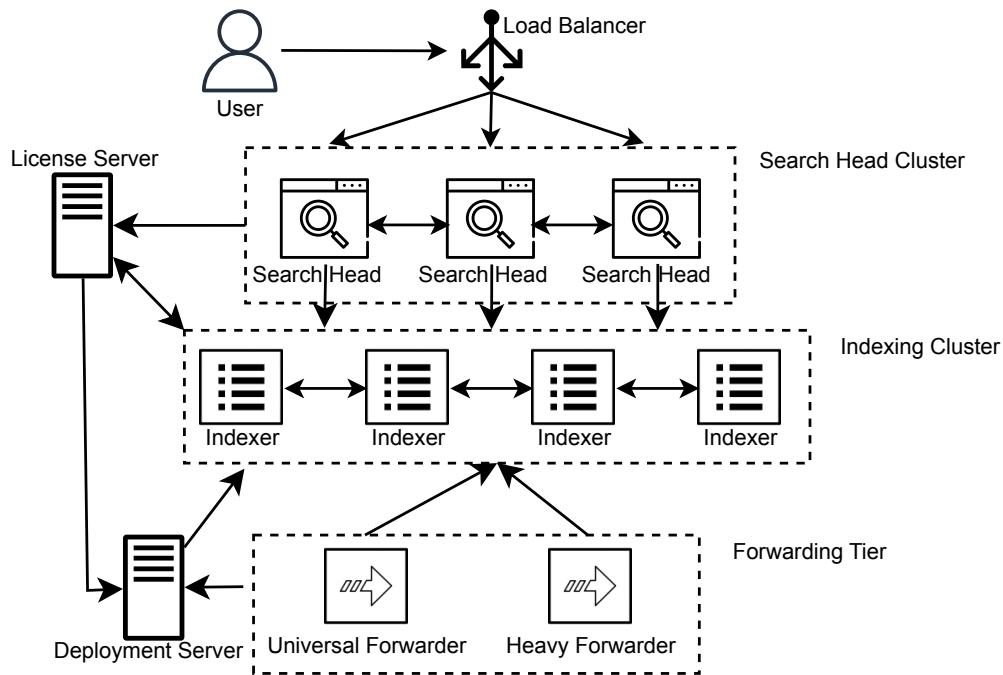
The service operating on the source device sends the selected events to the log collector component of the SIEM system, which proceeds to convert them into a unified, common structure by parsing and normalizing the inputs. The event is then fed into the correlation engine which analyzes and correlates the standardized information and proceeds to store it in an immutable way. The data is then extracted either via user defined searches or through continuous monitoring, which allows the detection of possible anomalies based on the current flows of events and their history.

The SIEM system produced by Splunk [11] is a globally known software used by major enterprises in all sectors, including telecommunications. For this reason, the project scope related on the investigation of SIEM systems will focus solely on it.

### 2.1.1 Splunk

Splunk was released in 2005 as a unified platform that assisted the IT department in troubleshooting technical issues by centralizing the storage of logs in a SIEM system. As the number of users grew, so did the features implemented, evolving into the product that is known as Splunk Enterprise Security [12]. As Splunk is also offering other cybersecurity products, the rest of the report will associate Splunk Enterprise Security with Splunk to improve the clarity of the text.

The current architecture of Splunk is composed of a three layer structure in which multiple entities cooperate to support its operations and provide services offered by a SIEM system. One layer is dedicated to interfacing with the devices generating logs, another focuses on the storage and indexing of such events, and the last is used to generate views of the information. Figure 2.2 shows the different elements and layers of the architecture and the interactions between them:



**Figure 2.2:** Simplification of the main components and communication paths of the Splunk architecture [13].

Starting from the bottom of the figure:

- The Forwarding Tier is mainly composed of two types of entities, both used to forward the flow of logged events to the indexers:
  - The Universal Forwarder consists of a simplified version of the Splunk software that is limited to solely performing the transmission of raw logs. This allows

it to not require many resources for its execution, with a limitation on the processing of information, which is delegated in its entirety to the indexers.

- The Heavy Forwarder consists of a full fledged version of the software with functionalities non relevant for the forwarding process disabled. Its main advantage is given by the ability to preprocess the logged data before transmitting it, reducing the load on the indexers with the limitation on the amount of resources needed for its operations.

Both types are directly provisioned and deployed on the machine the SoC has an interest in analyzing its logs and the Universal Forwarder is generally preferred to be used when possible.

- The Deployment Server provides a centralized way for the deployment of provisioning updates to the different forwarding agents.
- The Indexer Cluster is composed of multiple cataloging entities, known as indexers, which process the incoming logs into universally interpretable structures, other than saving the original data in an immutable format. A single forwarding element can send data to multiple indexers to optimize the ingestion times. The indexers are also exchanging information with one another to introduce redundancy and reduce the execution times of searches.
- The License Server monitors the status of all the different components and instances and checks if the license agreement is being respected.
- The Search Head Cluster groups multiple Search heads, which are used to extract information from the indexers and generate views for the users. All instances contain the same configuration, allowing scalability and distributing the load generated by the user actions.
- The Load Balancer interfaces with the user requests and assigns them to different search heads in order to optimize the execution of the tasks.

Splunk is also allowing the customization of its platform by enabling the installation of prebuilt or custom made applications which are suited to specific needs such as data analysis or monitoring [13]. This simplifies the work of the SoC by accessing already existing dashboards and tools without requiring manual configurations.

### **Search Processing Language**

The language used to perform searches over the logged information is called Search Processing Language (SPL), and instructs the search engine of Splunk to access specific events in the indexer and perform elaborations on them [12]. Each search consists of a string characterized by a syntax similar in format consisting of a combination of the Search

Query Language (SQL) and Unix commands. An example of the SPL syntax is shown in listing 2.1:

```
index=GameData  
| top user_country 5
```

**Listing 2.1:** SPL syntax example.

This search retrieves the information contained in the index *GameData* and outputs the five most common values of *user\_country* by using the search command *top*. The elaborations of the data are passed between different commands through the pipe operator (*|*).

Splunk also offers report generation capabilities, allowing users to visualize the results of a search using charts or panels through interactive dashboards. This enables the sharing of information to different stakeholders and their storing in a centralized place. Additionally, Splunk offers the possibility of generating alerts based on the results provided by a search. This happens when a correlation search, i.e. a search that runs periodically, begins to satisfy a series of conditions. When each of them is occurring, Splunk can trigger the execution of an automated action such as generating an alert on its system, but also send an email to specific addresses or run a particular script [13].

### Input Sources and License Costs

Splunk, and generally all proprietary SIEM systems, typically come at a cost for their usage. In the case of Splunk, its licensing plans are either based on: the ingestion volumes of log data per month, the number of entities forwarding the logs, i.e. the number of forwarders, or the computational resources needed to execute its operations and searches [14]. Therefore, the SoC has to define a tradeoff between the licensing costs and the quality or quantity of information. The same considerations can also be applied to the costs related to maintaining the information stored in Splunk. Independently its location, either on premises or on the cloud, a limitation on the storage space is present. It is therefore important to carefully identify which logs are considered to be the most valuable in terms of information contents to avoid large licensing or infrastructural costs.

## 2.2 Threats over the Telecommunication Infrastructure

According to the 2024 report of the Danish Center For Cyber Security (CFCS) related to cyber threats in Denmark [15], there are high risks for the country of becoming the target of attacks due to the current geopolitical situation. Based on threat landscape reports for the telecommunication sector released by the European Union Agency for Cybersecurity (ENISA) for 2024 [16] and the Global System for Mobile Communications Association (GSMA) for 2025 [17], the sector was targeted by the following threats:

1. Data Breaches and Leakages: caused by attacks leveraging vulnerabilities in systems misconfigurations, but also the reliance on third parties for the management of parts of the infrastructure, and increased complexity in the range of technologies used.
2. Denial of Service (DoS) and Distributed DOS (DDoS) attacks: aimed at the disruption of the communication services and mostly carried out using botnets of insecure Internet Of Things (IoT) devices, but also physical sabotages carried out in the physical infrastructure.
3. Malware: mainly targeting IT networks via email attachments. The core infrastructure is less subject to attacks, as it is not typically directly connected to the Internet,
4. Ransomware: caused by email attachments and targeting Linux software, typically used for the management of the core infrastructure.
5. Social Engineering: mainly performed through phishing tactics using compromised email addresses or voice calls.
6. Supply Chain Attacks: performed both in terms of hardware backdoors, but also focusing on open source software packages.

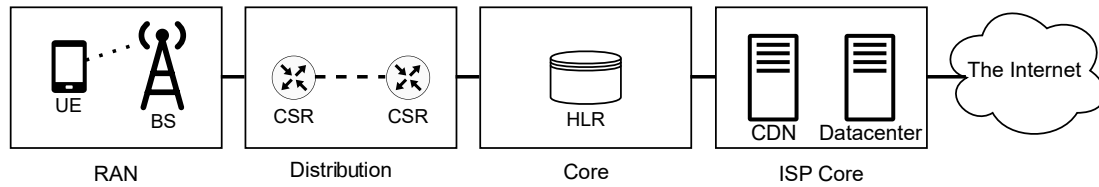
A more sector specific attack related to cellular networks can be found in the abuse of signaling systems, i.e. the protocols used to coordinate interactions between devices in telecommunication networks. In their security assessment reports, ENISA [18] and GSMA [19], underline how attacks on SS7 protocol are present and cannot be effectively stopped due to its working principles. Even though guidelines to mitigate the threats have been adopted by network operators, such as monitoring of the incoming signals or filtering specific message types, it is still an abused attack vector [18].

The following section 2.3, will introduce the infrastructure of a telecommunication network and section 2.4 will then focus on the SS7 protocol stack and its known attack vectors.

## 2.3 Infrastructure of a Telecommunication Network

Since the 4G and 5G networks have been designed to integrate cellular networks with the Internet, as introduced in chapter 1, the infrastructure of a telecommunication company evolved to accommodate those needs. This section will provide an overall view focused on the cellular networks sections and their elements, along with a reference to the overall architecture of an Internet Service Provider (ISP).

The main elements composing the cellular network infrastructure are found in: core networks, distribution networks, and mobile networks. Figure 2.3 showcases a uniform division of the network structure among the different cellular generations, as each one introduces different acronyms and new entities in the infrastructure.



**Figure 2.3:** Architecture of a Telecommunication network with a focus on mobile networks.

Starting from the left:

- The mobile network, also known as Radio Access Network (RAN) is responsible for providing a coverage area in which subscribers can connect to the cellular network of their network operator. Its main components are the User Equipment (UE), and the Base Station (BS).
- The Distribution network interconnects the mobile and core networks and it is used to provide connectivity to the BSs using high speed technologies such as fiber connections. The different RANs are interconnected to the network via a series of Cell Switched Routers (CSR), used for routing the data between the sections and provide synchronization [20].
- The core network is the backbone of the entire infrastructure. It provides the services and hosts the entities responsible for the management of the interactions between UEs. It enables the cellular networks to perform voice calls, send SMSs, and access the Internet. Each mobile generation defines its own entities to govern the aforementioned processes and hosts elements to have retro compatibility with previous generations, thus enabling connectivity between subscribers using different network generation technologies. A common entity between all generations, even though it may have a different name, is the Home Location Register (HLR). It is the central point of knowledge that stores information about all the subscribers managed by an operator.
- The ISP core network instead, is used by the core network to provide Internet services to the UEs. They can access the resources present in datacenters or Content Delivery Networks (CDN). If the resource cannot be found in the internal network, the request is sent to the Internet by connecting with networks managed by other ISPs [21].

As SS7 is a signaling protocol mainly used for the coordination in 2G networks, also known as Global System for Mobile Communications (GSM), the subsequent theoretical background will therefore focus on analyzing the GSM infrastructure with greater detail.

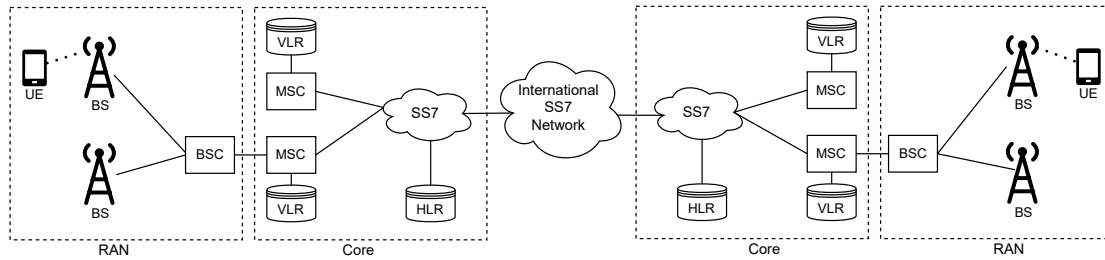
### 2.3.1 Elements of the GSM Infrastructure

Multiple network operators providing GSM coverage over a specific geographical area can coexist. Each of them is distinguished from the others via a Public Land Mobile Network



(PLMN) ID that identifies both the country it operates in and the actual operator code [22]. These operators are interconnected to allow the interaction between the UEs.

Figure 2.4 provides a graphical depiction of the entities for the Core and RAN networks, focusing on the use of SS7 and the relation between two PLMNs.



**Figure 2.4:** The GSM infrastructure and its elements with a focus on the signaling protocol [23] [24] [22].

- UE: are composed of a combination of a cellular phone and a Subscriber Identity Module (SIM) card that allows access to the cellular network of the PLMN that released it. Each SIM card internally stores, among others, the following information:
  - International Mobile Subscriber Identity (IMSI) number, used to uniquely identify a user in the global cellular network.
  - Authentication keys, used to authenticate the SIM card of a customer in the network.
  - Mobile Station International Subscriber Directory Number (MSISDN), consisting of the phone number associated with the SIM card.

These elements are universal to each generation of mobile networks. Each phone is also internally storing an International Mobile Equipment Identity (IMEI) number that uniquely identifies the physical device, which is passed along with the IMSI when interacting with the BS, thus the network itself. [22].

- The Base Station Controller (BSC) is the controller of one or more BSs. It is responsible for managing the exchange of radio communications through the BSs, and interfacing the user traffic with the Mobile Switching Center (MSC) [22].
- The MSC is responsible for managing the connection of a UE and supporting the handover procedures between BS, making it the core element for coordinating the interactions between UEs and the network. This implies interacting with different entities such as other MSCs to set up calls or the Short Message Service Center (SMSC) to send SMSs [24].

- The HLR is a centralized database that stores, in a permanent way, the information related to all the subscribers of a PLMN. It contains information used for the authentication of each SIM card issued by the operator, as well as the IMSI and MSISDN numbers to identify a SIM card, as well as the locations in terms of network area in which the subscriber is connected to the network [24]. The HLR is therefore used as the central source to verify the identity of a user before granting them access to the network as well as keeping track of the MSC that is handling its operations, allowing calls and SMSs to be forwarded to it.
- The Visitor Location Register (VLR) consists of a smaller, version of the HLR containing a temporary copy of the information related to the subscribers currently associated with a particular MSC. Instead of storing the network area, i.e. the VLR number associated with a subscriber, as the HLR does, it keeps track of ID of the BS providing coverage to the UE, facilitating the handover operations when a device physically moves and reducing the load on the HLR [22]
- The Equipment Identity Register (EIR) stores the IMEI information of the UEs, determining which devices are to be granted access and which should be restricted, i.e. when a device is reported as stolen, it should not be granted access to the network.

The protocol stack used for the coordination and communication between the different entities of a network, but also across different PLMNs, is SS7. The following section will provide an explanation of its working principles and later present its attack vectors.

## 2.4 Signaling System No.7

The SS7 protocol stack is used for the establishment and governance of all coordination mechanisms in GSM networks, ranging from device registration to the establishment of voice calls and forwarding of SMSs [25]. Despite its antiquity, as it was established in the 1980s, and being replaced with more advanced signaling protocols such as Diameter [26] for 4G networks, SS7 and its known vulnerabilities remain an attack vector at a global scale as long as GSM networks are supported. In fact, when a network operator supports 2G, it has to make sure its successors 4G and 5G maintain a retro compatibility layer that interfaces with it. This allows the interactions between UEs but at the same time ensures the presence of vulnerabilities that can be exploited even on the latest generations [27].

The majority of the operators in various countries are either planning to decommission 2G networks, Norway is an example [28], or are actively considering it. In Denmark, none of the operators were, in 2024, considering it [29]. Many low bandwidth communication systems such as IoT devices and sensors, as well as Machine to Machine systems like windmills, and emergency systems for elevators or cars are relying on 2G communications, implying that the decommissioning process could take up to a decade. It is also important

to mention that the SS7 protocol stack is still being used today to perform voice calls unless the PLMN has enabled voice exchanges over 4G, but also when a subscriber is roaming internationally [30].

### 2.4.1 Network Stack

The SS7 protocol stack consists of a series of layered protocols, in which each layer makes use of the abstractions provided by the previous level to provide services that are offered to the next [25]. Figure 2.5 makes a comparison between SS7 and the Open System Interconnection (OSI) stack, displaying the equivalence in functions between the OSI layers and the SS7 protocols.

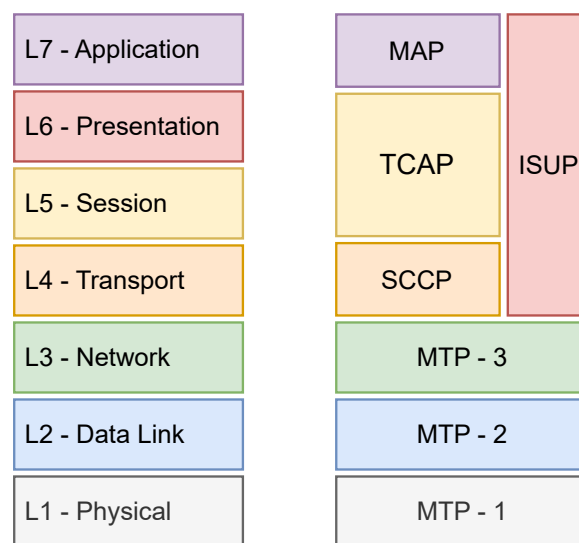


Figure 2.5: Comparison between the SS7 Protocol Stack and the OSI Layers [24].

Starting from the bottom:

- The Message Transfer Part 1 (MTP-1) protocol controls the way physical signals are modulated in order to be transmitted over the transmission medium, performing the equivalent functions of the Physical Layer in OSI.
- The Message Transfer Part 2 (MTP-2) protocol, divides the information in packets and performs flow control over the transmission between two devices, performing functions equivalent to the Data Link Layer.
- The Message Transfer Part 3 (MTP-3) protocol, performs the routing of messages, making use of the source and destination codes that uniquely identify them, making it equivalent to the Network Layer.

- The Signaling Connection Control Part (SCCP), is used to enhance the operations of MTP-3 and assign a unique address called Global Title (GT) to each Signaling Point (SP). Each entity operating in the GSM architecture, described in subsection 2.3.1, apart from the UEs and BSs is assigned a GT that identifies them in a PLMN, thus the global GSM infrastructure.

Three main protocols are used to define the interactions across the SS7 network [25]:

- The Integrated Services Digital Network User Part (ISUP) protocol is used for coordinating the establishment of the circuit switched connection between the origin MSC and the destination one, i.e. establishing, maintaining and terminating a call.
- The Transaction Capabilities Application Part (TCAP) protocol allows the exchange of signals that are not related to the establishment of circuits between the different SP. This allows the different entities in the GSM infrastructure to concurrently exchange messages. Its functions are then used by the MAP protocol to perform operations such as interfacing with the HLR.
- The MAP protocol is used to coordinate the communication between the MSCs and HLRs across the GSM infrastructure. It enables the coordination and management of the subscriber thus allowing the operations performed by the HLR.

The protocols building on top of MTP-3 are the ones enabling the different network functionalities of the GSM infrastructure. They are therefore the most exploited protocols to perform attacks on the 2G networks. The following subsection will introduce some attack examples for the SCCP, ISUP, TCAP, and MAP protocols.

### 2.4.2 Attack Vectors

As the SS7 suite was developed in the 1980s, when only a handful of PLMNs were operating over a close network, security was not the main concern as trust between different network providers was considered implicit. However, as more countries and new operators implemented their cellular infrastructure, and new network generations integrated mobile networks with the Internet, as previously discussed in chapter 1, trust was no longer enough to guarantee the security of the global infrastructure [24]. As virtually no security mechanisms are present in the stack itself, the presence of malicious or compromised operators can lead to attacks targeting all the subscribers in the GSM network.

Once an attacker has gained access to the SS7 infrastructure, multiple attack vectors can be abused. The network attack taxonomy made by GSMA [31], as well as Ullah et. al [24], provide a comprehensive view of them, both from a general view, but also with a focus on the actions the network operator can do to protect itself. The following list provides a generalization of the known attack possibilities based on the protocol that can be exploited:

- SCCP
  - Network Scans: SCCP packets can be used to scan the SP infrastructure of a PLMN network to obtain knowledge about the active GTs. This allows the attacker to map the network entities in preparation for targeted attacks.
  - Impersonation of Network Elements: An attacker spoofs its identity with the one of a SP by using its GT the communication, making it therefore able to intercept the traffic it is redirecting.
- ISUP
  - Caller ID Spoofing: An attacker manipulates the calling number in the packet to another MSISDN, allowing it to hide its real identity.
- TCAP
  - TCAP Scans: An attacker sends specific probe packets to known GTs to retrieve information such as the protocol version and the available services, performing the reconnaissance for vulnerable GTs.
- MAP
  - Location Tracking: Multiple MAP messages can be exploited in order to obtain the GT of the BS providing coverage to the target. It can be performed by abusing messages such as call setup or SMS forward requests. This attack can be performed both if the device is connected to its home network or if the device is roaming.
  - HLR Lookup: the attacker is able to retrieve the IMSI number of a SIM card associated with a customer by interrogating the HLR managing the identity of the target.
  - DoS: An attacker can block a subscriber from communicating with the cellular network by performing malicious location updates on the HLR or deletions from the VLR providing coverage to a specific IMSI number.
  - Calls and SMS interception: An attacker impersonates or compromises an MSC, as well as maliciously swapping its identity with the target in the HLR to redirect the contents of a call or a message to itself.

Having considered that multiple attack types can be performed in the SS7 stack, the scope of the project will focus on the ones directly targeting the users. As the majority of them are found in abuses of the MAP protocol, the following subsection will analyze particular MAP attacks with greater detail on the invoked procedures and communication phases between the SPs in the network.

### 2.4.3 Attacks over MAP

MAP, thanks to the services offered by TCAP, makes use of service calls to allow the interaction between the different entities in the infrastructure, from the UE to HLRs and MSCs. However, as no verification of the identity of the calling SP is present by default, if no security hardening practices are implemented, an attacker could potentially target any user in the world. The following subsections will analyze attacks related to HLR Lookup, Location Tracking and SMS and Call Interceptions.

#### HLR Lookup and Location Tracking

In these scenarios, the attacker is able to gain knowledge about the identity and location of the target by either interacting with the HLR or directly with the MSC/VLR managing the subscriber [31]. Both attack modes treated in this subsection enable the retrieval of the location of the target but differ in the way they are carried out.

#### Any Time Interrogation

The Any Time Interrogation (ATI) message type allows, as the name might suggest, an interrogation request to the HLR about one of its subscribers that can be executed at any time. Even though this type of message can be legitimately used by network operators to obtain knowledge of the status and location of a subscriber, it can be easily exploited for malicious purposes, as no reason is required to invoke it. Figure 2.6 describes the flow of messages between the attacker and the HLR and MSC providing network coverage to the target:

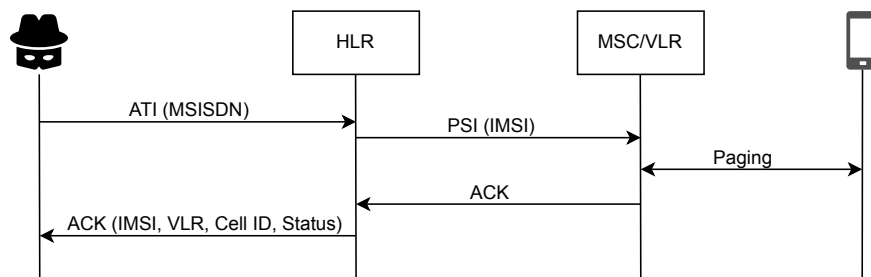
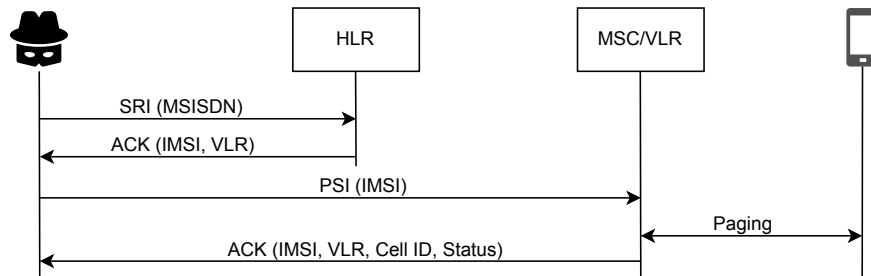


Figure 2.6: Attack flow of the ATI attack [24].

The attacker, impersonating the HLR of a different network operator, sends an ATI request including the MSISDN of the target to the HLR managing its identity. The HLR, after consulting its contents, contacts the MSC handling its operations with a Provide Subscriber Info (PSI) message requesting information for the specific IMSI associated with the MSISDN. The MSC proceeds to page target through the BS it is connected to, based on the information its VLR contains. Once an answer is received, the MSC informs the HLR about its status and location, which is sent back to the attacker.

### Send Routing Info

The Send Routing Info (SRI) message allows a network operator to obtain from the HLR the location, i.e. the identity, of the MSC handling a subscriber [25]. This is done in order to establish a call or forward an SMS to the target. However, as SS7 does not have security or identity verification mechanisms, once the attacker has knowledge about the MSC to contact, it can proceed to locate a user instead of continuing further with the call or SMS process. Figure 2.7 showcases this process:

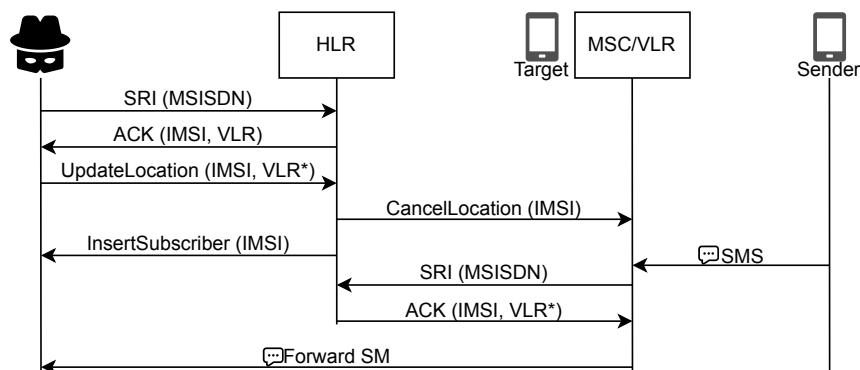


**Figure 2.7:** Location gathering attack flow based on the SRI message [24].

The attacker, impersonating an MSC, requests the HLR of the target to provide the routing information, either for a call or SMS, through a SRI in which it indicates the MSISDN. The HLR answers with the IMSI and location of the MSC. The attacker now proceeds to send a PSI request of a particular IMSI to the MSC of the target instead of continuing the usual flow of messages. The MSC returns the updated location and status after paging the target.

### SMS interception

In this scenario, the attacker impersonates the MSC providing service to the target and maliciously updates its location in the HLR. This allows the attacker to reroute the flow of calls and messages destined for the target, as the HLR believes it is being served by the MSC impersonated by the attacker, causing the target to not receive such events on its UE. Figure 2.8 showcases the attack flow of an SMS interception attack:



**Figure 2.8:** SMS interception flow.

The attacker, acting as a MSC, retrieves the IMSI of the target by beginning the procedure to send an SMS through the SRI message for SMSs. Once it receives an answer from the HLR of the target, it informs the latter of the presence of a new subscriber, to its network area by sending a `UpdateLocation` containing the IMSI of the target and the location of its VLR. The HLR updates its memory and informs the actual MSC serving the target to delete the IMSI of the subscriber on its records with a `CancelLocation`. The HLR then informs the MSC of the subscriber to add the IMSI of the target to its VLR. Once another subscriber sends an SMS to the target, the HLR provides the forwarding location to the MSC of the attacker, redirecting the SMS.

After having gained knowledge about Splunk, the network infrastructure of telecommunication operators with a focus on GSM networks, and attacks over the SS7 protocol stack with a focus on MAP, the final problem statement is presented:

*How to define searches in Splunk to detect the presence of SS7 attacks over MAP in the logs produced by the core network infrastructure, while maintaining acceptable detection levels with a limited amount of input sources?*



## Chapter 3

# Detection of Attacks

*This Chapter begins with section 3.1 by discussing the problems faced related to the access of network logs from the core infrastructure and addressing possible solutions to solve them. Subsequently, section 3.2 discusses how the Java SS7 (JSS7) simulator has been used to generate logs of network attacks along with the issues faced in ingesting them into Splunk. The identification of the attack indicators along with the definition of searches that can detect them is discussed in section 3.3.*

### 3.1 Issues Faced and Possible Solutions

The possibility of performing the detection of SS7 attacks in the cellular infrastructure implies the need to have access to the network logs produced by the various SPs. As these entities are handling sensitive data related to the call logs, SMS messages, and the location of each subscriber in a country, working with real world data will be a problem. This is due to the fact that having access to such information would be a clear violation of privacy and the European General Data Protection Regulation (GDPR) regulation [32]. The following subsections will first analyze possible solutions that could be performed to access real data and then focus on the use of simulators.

#### 3.1.1 Compliance with Regulations

As the thesis work has been carried out in Denmark, the GDPR legislation protecting the privacy of a person must be respected. This prohibits the access to personally identifiable information without the previous explicit consent of its owner. However, according to the GDPR, access to such data could be granted on the grounds of consent or legitimate interest:

- Regarding the ground of consent, contacting the customers of the network operator and asking for their explicit consent to analyze their network traffic for research purposes could be a possibility. However, even if this authorization could be granted

for an extremely limited amount of time, giving access to such personal information would be seen by virtually every person as a violation of privacy, resulting in a very limited number of participants.

- Regarding the legitimate interest, the network operator could advertise an opt in security protection package against network attacks targeting their customers. By promoting a defense against location exposure or unlawful interception attacks in exchange for the network traffic associated with a user, a greater number of participants would be prone to giving consent. This is because a customer would see a benefit in exchange for their data, other than providing an economical reward if such protection comes at a cost.

Both ideas could be implemented with the appropriate planning and advertisements. However, the execution times and coordination with different departments would require an amount of time and preparation that exceeds the length of the semester allocated for the thesis project.

It is also worth mentioning that, as cellular networks are part of the critical infrastructure of a country, it is necessary to comply with national laws and internal company regulations. In order to be granted access to the data related to both network traffic and actual infrastructural implementation, it is necessary to be granted an appropriate privileged access level and receive security training.

As working with both raw data from consenting customers, and requesting to be granted privileged access to such information exceeds the time allocated for the thesis project, the access to real network traffic is therefore excluded from the scope of the report. However, one might argue that making use of anonymization techniques could allow the thesis work to be carried out on non identifiable data coming from a real system. The following subsection will therefore take into consideration possible approaches that could be implemented.

### 3.1.2 Data Anonymization

Majeed et al. [33] define multiple ways to achieve data anonymization in order to comply with both regulations and preserve the privacy of users. As seen in subsection 2.4.2, the SS7 traffic contains information used to identify both the SPs of the GSM infrastructure, but also the identity of a subscriber. These identifiers can be classified into two different groups called Direct Identifiers and Quasi Identifiers. A Direct Identifier is associated with the information that allows to identify a person, in a unique way, IMSI and MSISDN in this case. A Quasi Identifier is related to the elements that can be used to indirectly identify the user, and the location expressed in terms of MSC GT and BS ID. If the aforementioned attribute types are obfuscated or removed from the real network logs, it might still be

possible to make use of actual signaling data for the thesis work. If supposing that both direct and indirect identifiers are represented as unique integer values, it would be possible to perform two possible actions on the data [33]:

- **Perturbation:** achieved by substituting the identifier with the output of a hashing algorithm. Given the sensitivity of the information, multiple runs of the same or a combination of different hashing functions could be used. As the information is still present in the perturbed values, a powerful brute force attack could still reveal the actual content of the information.
- **Anatomization:** obtained by mapping all the identifiers in a different table and substituting the values with the ID associated with the position in the latter. If the table is not disclosed, the actual information is not included in the data.

As an already existing solution performing the anonymization of real network traffic logs was not present and considered that its eventual development or manual anonymization would require an unplanned use of internal resources, the access to anonymized data from a real network source is excluded from the scope of the thesis. It is also worth mentioning that, even though real data could be used, such logs might not contain SS7 network attacks, hindering the project work.

Therefore, the identified viable solution to access network logs containing SS7 attacks can be found in making use of a simulator or a freely accessible dataset of a network log. By using simulated information, no personal information is present in them, thus not requiring the necessity of compliance with privacy regulations.

### 3.1.3 Simulators and Datasets

Two different open source simulators have been taken into consideration for the generation of the network logs containing the SS7 attacks: JSS7 [34] and SigPloit [35]. The former has been developed over a fork of the SS7 simulator made by RestComm [36] and allows running the entire network simulation over a single machine and already includes attack examples. The latter consists of a more advanced simulator that requires the use of a dedicated virtual testing scenario or direct access to the SS7 network infrastructure. As no testing environment was already being used in Telenor, and considering the required time and resources needed for its development, SigPloit was excluded from the scope of the project. The direct access to the SS7 infrastructure is, for the reasons expressed in subsection 3.1.1, not possible.

The JSS7 simulator has also been used by Guo et al. [37] to generate a dataset [38] containing network information expressed in a format similar to real data provided by a Luxembourgian telecommunication provider. As the research carried out by Guo was centered around a Machine Learning approach for the detection of SS7 attacks, its contents

contained features such as counters and time indicators between messages, making it not a suitable source for the detection of attacks based on the contents of MAP messages. As no other already existing publicly available datasets have been found, JSS7 has therefore been used as the simulator to generate network logs of SS7 traffic. The following subsection will provide a technical overview of the scenarios that have been simulated.

## 3.2 Generation and Import of Network Logs

The JSS7 simulator allows the generation of a scenario in which up to three PLMNs, each consisting of their own HLR, MSCs and UEs, are interacting with each other. The traffic generated by the entities has been captured in a *.pcap* file using Wireshark [39] set to listen on the *loopback* network interface of the machine running the experiments. The requirements for running the simulator are found in using a Linux machine with Java 8, Wireshark and the Stream Control Transmission Protocol (SCTP) tools installed on it. The simulator abstracts the MTP-1 and MTP-2 layers, building on top of SCTP the remaining SS7 suite, going from MTP-3 to SCCP, TCAP and ultimately MAP.

The types of simulations run by JSS7 are divided into two different modes: attack demonstrations and a realistic scenario. The former produces three different scenarios consisting of isolated attacks over the MAP protocol related to Location tracking and SMS Interception. The latter consists of a more realistic simulation in which multiple UEs, other than the target, interact with each other, thus adding noise to the network logs by mimicking its normal behavior. The attacks included in the demonstrations are carried out in the complex at randomly determined times. The following simulations have been generated:

- Location tracking using ATI
- Location tracking leveraging SRI
- SMS Interception
- Complex Simulation

The first three consist of a demonstration scenario composed of two separate PLMNs in which only one subscriber is acting as a target, and the attacker operates from a different network by impersonating its HLR or MSC. The remaining one, the complex simulation, consists of a setting composed of three different networks with a subscriber registered in each of them. Due to some misconfigurations in the simulation code, the attacker and target appear to be placed in the same PLMN. As the simulator consists of a binary file invoked by a shell script, it has not been possible to solve such issue.

The work carried out throughout the report will, due to timing concerns, focus on the identification of the location tracking attacks and their detection in the complex simulation,

excluding the SMS interception from the scope of the report. The following subsection will focus on the problems faced with importing *.pcap* files in Splunk and section 3.3 will focus on the detection of the attacks.

### 3.2.1 Import of Data in Splunk

Even though Splunk has the ability to extract information from the *.pcap* files generated by Wireshark by making use of a specific application that can be installed on it, it has been necessary to reach a compromise in the representation format used for the ingestion of the information. The problem seemed to be related to the fact that the entire insertion procedure consisted of a one time ingestion and not a continuous flow, causing problems in its elaboration. Splunk is in fact optimized for the ingestion of data inputs from a forwarder rather than one time insertion. When the upload of the *.pcap* files have been attempted, the only information that was possible to extract was related to the high level overview of the traffic data, i.e. time, source address, destination address, and textual info of the packet. This meant that the actual information contained in the SS7 packets was not going to be present in the Splunk Indexers, limiting the thesis work.

Wireshark also offers different formats in which the data can be structured before it is saved: JavaScript Object Notation (JSON), text, Comma-Separated Values (CSV), C array or Extensible Markup Language (XML) formats:

- JSON format: Splunk had difficulties in performing the automatic separation of each packet into a single event, but was also not able to extract their fields and respective values from the structure. It was also not possible to use Regular Expressions to perform such separation, as the format used to delimit each of them was also used to separate the single protocol headers in each event.
- Text and CSV formats: As their contents only saved the high level overview, it did not allow the ingestion of the full information of a packet, exactly as in the case of the direct import of the *.pcap* file.
- C array format: As it consists of a binary representation of the contents, it is unintelligible.

Regarding the XML format, Wireshark makes use of the Packet Description Markup Language (PDML) file format, structuring the single packets and their content in an XML format. Splunk was able to separate the single events without major problems, making it the only feasible option to ingest the simulation logs. Figure 3.1 shows a snippet containing the contents of an event limited to a part of the SCTP header in Splunk.

```

<proto name="sctp" showname="Stream Control Transmission Protocol, Src Port: 8019 (8019), Dst Port: 8020 (8020)"
size="28" pos="34">
  <field name="sctp.srcport" showname="Source port: 8019" size="2" pos="34" show="8019" value="1f53"/>
  <field name="sctp.dstport" showname="Destination port: 8020" size="2" pos="36" show="8020" value="1f54"/>
  <field name="sctp.verification_tag" showname="Verification tag: 0x094fd588" size="4" pos="38" show="0x094fd58
8" value="094fd588"/>
  <field name="sctp.assoc_index" showname="Association index: disabled (enable in preferences)" size="0" pos="3
4" show="65535"/>
  <field name="sctp.port" showname="Port: 8019" hide="yes" size="2" pos="34" show="8019" value="1f53"/>
  <field name="sctp.port" showname="Port: 8020" hide="yes" size="2" pos="36" show="8020" value="1f54"/>
  <field name="sctp.checksum" showname="Checksum: 0x00000000 [unverified]" size="4" pos="42" show="0x00000000" v
alue="00000000"/>
  <field name="sctp.checksum.status" showname="Checksum Status: Unverified" size="0" pos="42" show="2"/>
  <field name="sctp.chunk" showname="SACK chunk (Cumulative TSN: 3181825453, a_rwnd: 106496, gaps: 0, duplicate
TSNs: 0)" size="16" pos="46" show="" value="">
    <field name="sctp.chunk_type" showname="Chunk type: SACK (3)" size="1" pos="46" show="3" value="03">
      <field name="sctp.chunk_bit_1" showname="0... .... = Bit: Stop processing of the packet" size="1" pos="46"
show="False" value="0" unmaskedvalue="03"/>
      <field name="sctp.chunk_bit_2" showname=".0.. .... = Bit: Do not report" size="1" pos="46" show="False" va
lue="0" unmaskedvalue="03"/>
    </field>
  </field>

```

**Figure 3.1:** Example of the representation format of a packet in the *.pdml* format ingested in Splunk.

Even though the extraction of the specific fields was not possible, no other formats containing the details of the packets were able to be ingested and accurately separated. For these reasons, the *.pdml* format was used as the base for the definition of the searches. This will pose a limitation on the human intelligibility factor, as it will be necessary to manually identify the attack indicators and extract them through specific commands. This tradeoff is however acceptable as it was the only way to have access to the information dump of the signals in Splunk.

Now that the simulation data has been ingested in Splunk, it is possible to start defining the searches. The next section will discuss the reasoning behind the identification of the attack patterns and their detection using the SPL.

### 3.3 Identification of Attack Indicators

Before being able to define the searches to be run in Splunk, it has been first deemed necessary to gain practical knowledge about the MAP messages used for the location tracking attacks using the ATI and RSI procedures. Therefore, the scenarios containing the isolated attacks have been used as a starting point to easily identify the dynamics of such attacks using the SPL, and later verify their accuracy in the realistic simulation.

Overall, each search has been developed in the optic of reducing the number of false negatives as much as possible. If an attack flow is not detected, the reliability of a search would decrease, as further investigations would not be performed. The same reasoning is

also applied to false positives, as large unrelated signaling flows would introduce noise in the result, impacting the accuracy of the search.

### 3.3.1 Any Time Interrogation

Regarding ATI attacks, GSMA states that ATI messages should be rejected by default when received by any SP outside of some trusted internal entities unless an agreement has been signed between two network operators due to their infrastructure implementation [40]. Assuming that a network operator has not signed any agreement related to the use of such message and that its invocation from an internal SP should be considered as an indication of malicious activity due to the existence of other less invasive messages to obtain the MSC serving a subscriber, such as RSI, the search should therefore focus on the detection of each single ATI message being received by an HLR.

By using Wireshark, it has been possible to observe the attack dynamics in the isolated log in a structured and easily intelligible format. This allowed a swift identification of the important packets and their contents, which fields will be used for the definition of the Splunk search. Figure 3.2 displays the two packets used to invoke the ATI request and the answer from the HLR:

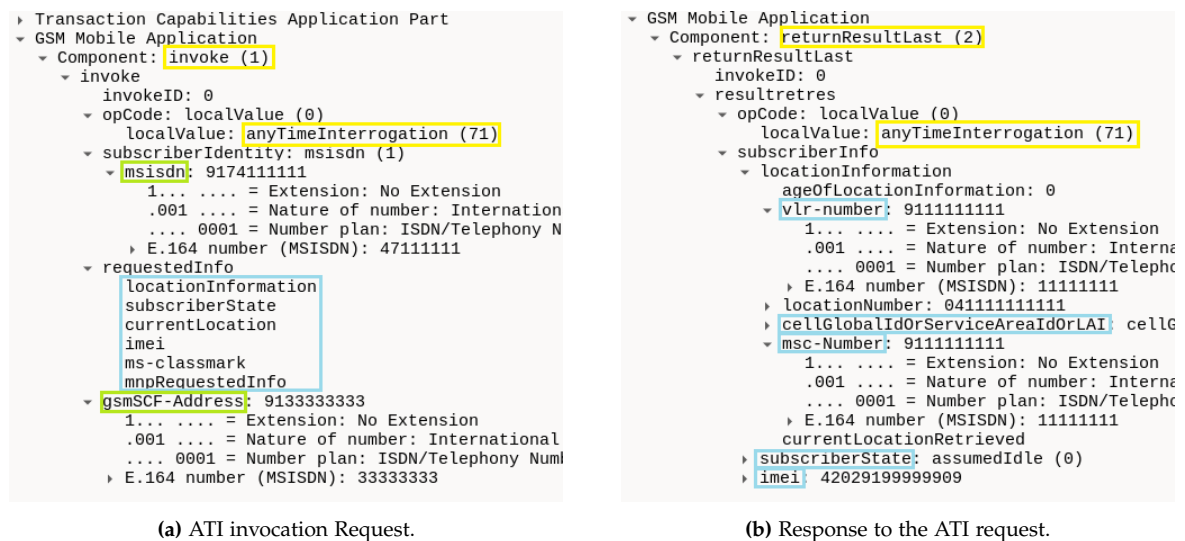


Figure 3.2: Flow of ATI attack seen from Wireshark.

For both figures, the fields in yellow identify the type of message, either request or response, and its category, specified as ATI. For Figure 3.2a, the ones in green, identify the MSISDN of the subscriber which should be queried, and the identity, specified as a GSM Service Control Function (gsmSCF), of the GT interacting with the HLR of the target. The fields in blue specify the kind of information that the HLR should be providing to the

attacker. For Figure 3.2b, the fields in blue show the information that the attacker was able to gather as a result of the ATI attack.

The correct detection and identification of an ATI attack shall therefore consider the following fields used within the MAP header:

- `localValue:anyTimeInterrogation (71)`
- `Component:invoke (1)`

### Definition of the ATI Search

By centering the search around the aforementioned fields, Splunk is able to select those events that consist of the invocation of an ATI. The translation of those needs in SPL is found in listing 3.1. Other than the aforementioned fields, the following search contains commands related to the extraction through regular expressions of the value of the fields related to the calling GT, the MAP request and the MSISDN of the target. This information will be later used in chapter 4 for the definition of the correlation search that can be periodically run and alert the security team of eventual detections.

```
index = main host=SS7TestData source="SS7 ATI.pdml"
| rex field=_raw "<field name=\"sccp\.calling\.digits\"[^\>]*
  value=\"(?<calling_gt>\d+)\ \""
| rex field=_raw "showname=\"localValue:\s*(?<action>.*?)\s?\(\\
  d+)\ \""
| rex field=_raw "<field name=\"gsm_map\.msisdn\"[^\>]*value
  =\"(?<msisdn>\d{10,15})\ \""
| search "Component: invoke (1)" "localValue:
  anyTimeInterrogation (71)"
```

**Listing 3.1:** Splunk search for the detection of an ATI attack.

The search instructs the search engine to access, via `index=<name>`, the main index, select the network log source containing the ATI attack, via `source=<fileName>`, from the data source identified by `host=<name>`. The retrieved event logs go through three different regular expressions aimed at extracting, in order, the field associated with the GT of the caller from the SCCP header, `calling_gt`, the MAP request, `action`, and the MSISDN of the target, `msisdn`. The resulting data is then passed to the `search` command, which finds in those logs that contain the invocation of a request, `Component:invoke (1)`, and the ATI MAP message `localValue:anyTimeInterrogation (71)`.

The resulting search has then been applied to the complex simulation by changing the source in which the events should be retrieved to its *.pdml* file to "SS7 Capture.pdml".



Figure 3.3 shows a visualization in a table format, which search can be found in listing A.1, of the obtained output from the complex simulation:

_time	calling_gt	action	msisdn
2025-04-23 17:51:48.331	11111111	anyTimeInterrogation	9164111111
2025-04-23 17:51:42.326	11111111	anyTimeInterrogation	9174111121
2025-04-23 17:51:26.175	11111111	anyTimeInterrogation	9174111111

Figure 3.3: Tabular representation of the ATI search applied to the realistic simulation.

A table shows the time of ingestion of the packet, the GT invoking the HLR request, the type of MAP action performed and the MSISDN of the target. The output resulted in 248 detections, all of which should be taken into consideration given the initial assumptions for the definition of the search.

### 3.3.2 Send Routing Info

The analysis through Wireshark of the network logs for the isolated SRI attack made it possible to determine that the message used to request the location is `SendRoutingInfoForSM`, which is used to obtain routing info to forward an SMS. Figure 3.4 showcases the SRI and PSI messages used in the attack:

```

Message Type: Unitdata (0x09)
... 0001 = Class: 0x1
0000 ... = Message handling: No special options
Pointer to first Mandatory Variable parameter: 3
Pointer to second Mandatory Variable parameter: 1
Pointer to third Mandatory Variable parameter: 21
> Called Party address (9 bytes)
> Calling Party address (9 bytes)
  > Address Indicator
    SubSystem Number: MSC (Mobile Switching Center)
    [Linked to TCAP, TCAP SSN linked to GSM_MAP]
  > Global Title 0x4 (7 bytes)
    Translation Type: 0x00 (0)
    0111 ... = Numbering Plan: ISDN/mobile (0x7)
    ... 0010 = Encoding Scheme: BCD, even number
    .000 0100 = Nature of Address Indicator: Inter
    > Calling Party Digits: 33333333
> Transaction Capabilities Application Part
> GSM Mobile Application
  > Component: invoke (1)
    > invoke
      invokeID: 0
      > opCode: localValue (0)
        localValue: sendRoutingInfoForSM (45)
      > msisdn: 9174111111
      sm-RP-PRI: True

```

(a) Invocation of the SRIforSM message.

```

Message Type: Unitdata (0x09)
... 0001 = Class: 0x1
0000 ... = Message handling: No special options
Pointer to first Mandatory Variable parameter: 3
Pointer to second Mandatory Variable parameter: 1
Pointer to third Mandatory Variable parameter: 21
> Called Party address (9 bytes)
> Calling Party address (9 bytes)
  > Address Indicator
    SubSystem Number: MSC (Mobile Switching Center)
    [Linked to TCAP, TCAP SSN linked to GSM_MAP]
  > Global Title 0x4 (7 bytes)
    Translation Type: 0x00 (0)
    0111 ... = Numbering Plan: ISDN/mobile (0x7)
    ... 0010 = Encoding Scheme: BCD, even number
    .000 0100 = Nature of Address Indicator: Inter
    > Calling Party Digits: 33333333
> Transaction Capabilities Application Part
> GSM Mobile Application
  > Component: invoke (1)
    > invoke
      invokeID: 0
      > opCode: localValue (0)
        localValue: provideSubscriberInfo (70)
      > IMSI: 24201111111110
      > [Association IMSI: 242011111111110]

```

(b) Invocation of the PSI message.

Figure 3.4: SRI and PSI messages used in the location tracking attack.

For both figures, the elements in yellow identify the invocation of a request and its type. The element in green showcases the GT of the calling party, i.e. the attacker. For Figure 3.4a, the MAP message type is `SendRoutingInfoForSM`, and the element in blue highlights the MSISDN of the target. Figure 3.4b is the PSI message that the attacker sends once the HLR of the target provides the identity of the MSC, and the element in blue showcases the transition from the MSISDN to the IMSI. The MSC serving the target will, as described in subsection 2.4.2, page it and provide an update on the location directly to the attacker, evading an eventual blockage put in place to prevent the location tracking using the ATI. Since the simulation in JSS7 has only been made using the HLR and MSC SPs, the latter was also used to represent the functionalities of the SMSC.

### Definition of the RSI Search

In an ideal scenario, the search should combine the knowledge of two fields present in three different packets:

- Caller ID from the *RSI* and *PSI* messages, allowing the identification of the GT of the attacker.
- IMSI from the *RSI response* and the invocation of the *PSI* messages, to identify the target.

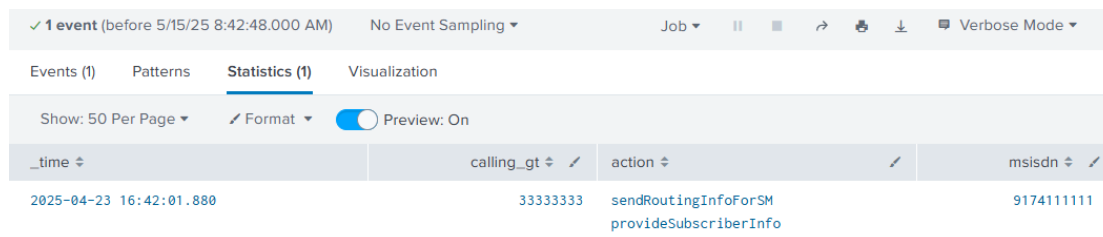
The combination of a PSI message for an IMSI obtained after the invocation of an RSI which provided as an answer such ID would be able to accurately correlate the events, thus detecting the attack. This reasoning can be successfully translated in a Splunk search by intersecting the three packets within a defined time span. However, due to time limitations, a simplified version merging only the invocation of the PSI and RSI messages performed by the same calling GT, within a certain time interval, that does not take into consideration the IMSI number has been realized. Listing 3.2 displays it, along with the extraction of the field used in chapter 4 for the creation of the correlation searches:

```
index = main host=SS7TestData source="SS7 PSI.pdml"
| rex field=_raw "name=\"sccp\.calling\.digits\".*?value=\"(?<calling_gt>\d+)\\"
| rex field=_raw "showname=\"localValue:\s*(?<action>.*?)\s?\"(\d+)\\"
| rex field=_raw "<field name=\"gsm_map\".(?:sm\\.)?msisdn\"[^>]*value=\"(?<msisdn>\d{10,15})\\"
| transaction calling_gt maxspan=1s
| search "localValue: sendRoutingInfoForSM (45)" AND "
  localValue: provideSubscriberInfo (70)" AND "Component:
  invoke (1)"
```

**Listing 3.2:** Splunk search for the detection of an RSI attack.

The search begins, as the previous one, by identifying the data source, the RSI network log in this case, its host and index storing it. Three regular expressions extract the GT of the attacker, the MAP actions performed and the MSISDN through rex. The command transaction is then merging in a single entity, all the events having the same calling GT that were generated within a time span of one second. The resulting events containing an invocation of an RSI and PSI are then selected to detect the attack.

The resulting search was also tested using the network logs of the complex simulators simulation. During the multiple iterations necessary to define the search in listing 3.2, the output ended up displaying two merged events which had no relations between each other as they were part of separate flows of communications. As only false positives were being detected, a manual check of the complex simulation logs was performed in Wireshark. It has been found that the only type of attack present in such file was related to the ATI attack, as the PSI messages have only been invoked after they received the ATI request from the attacker. Therefore, Figure 3.5 showcases, in a tabular view, the resulting output of the search performed for the overall host, i.e. all the simulation logs discussed in section 3.2, were taken into consideration:



_time	calling_gt	action	msisdn
2025-04-23 16:42:01.880	33333333	sendRoutingInfoForSM provideSubscriberInfo	9174111111

**Figure 3.5:** RSI search applied to all the simulated logs. The search command that produced this result can be found at listing A.2.

The table showcases the time in which the SRI message was logged, the GT of the attacker, the MAP invocations and the targeted MSISDN. As no other RSI attack was present in the logs, only one event has been detected.

This search can be performed with the assumption that the entirety of a PLMN core network is able to log its network traffic, meaning that all SPs including the MSCs should be included. A further consideration on the aspects related to the costs and ingestion volume will be discussed in section 5.1. The following chapter will complete the definition of the searches by exploring how the detection of attacks can be automatized.

## Chapter 4

# Creation of the Correlation Searches

*This chapter considers the automatization of the execution of the searches that have been defined in the previous chapter through the use of correlation searches. The information needed for their definition is found in section 4.1, the association with different cybersecurity frameworks is handled in section 4.2, the timing aspect is discussed in section 4.3, and the actions taken when an attack is detected are elaborated in section 4.4.*

After having identified and translated the attack patterns into two searches, the RSI and ATI attacks can be detected using Splunk. However, this can only be performed upon manual execution. To automatize the process, it is possible to use the correlation search mechanisms, which enables the Active Response characteristic of SIEM systems, as discussed in section 2.1. It is important to mention that the actions taken in case of detection do not remediate or prevent future attacks, a behavior typical of a Security Orchestration, Automation, and Response (SOAR) system, but serve as a notification mechanism to further investigate the matter.

In order to configure a correlation search, it is necessary to provide information related to the following aspects:

- The *definition* of the search itself and the context in which it should be run.
- The *mapping* with the security controls of different cybersecurity frameworks.
- The *timing* in which the search should be executed.
- The *adaptive response actions* taken by Splunk in case of detection.

The following subsections will therefore analyze the configuration for each of the defined searches.

## 4.1 Definition

Regarding the definition of a correlation search, Splunk defines multiple fields to be filled out. Other than the *search* itself, it is necessary to associate it with a *name* that is used to identify the correlation search, as well as a *description* explaining the actions and the type of information being processed. The *application* and its *application context* also need to be specified, identifying which functionality offered by Splunk should be used, as well as the type of information that can be processed. A field called *mode* is also present and specifies if the search should be defined on the go through a guided menu or if it has already been defined in the SPL.

Regarding the ATI attack, the *name* given to the correlation search has been defined with the optic of identifying the country of the network operator in which it was created and the ATI attack itself, resulting in Tndk - SS7 - AnyTymeInterrogation Attack Detection. The *application* and *application context* are both specified to Enterprise Security, as it is the SIEM system developed by Splunk and the information to be retrieved has been ingested in it. The *description* specifies the actions performed by the search, the *mode* is set to manual as the search has already been defined, and the actual *search* is the one present in listing 3.1. Figure 4.1 showcases the screenshot of such configuration:

The screenshot shows the 'Correlation Search' configuration page in Splunk. The fields are filled as follows:

- Search Name:** Tndk - SS7 - AnyTymeInterrogation Attack Detection
- App:** Enterprise Security
- App Context:** Enterprise Security
- Description:** This correlation search detects AnyTymeInterrogation message requests sent to the HLR
- Mode:** Manual (selected over Guided)
- Search:**

```
index = main host=SS7TestData source="SS7 ATI.pdml"
| rex field=_raw "<field name=\"sccp\calling\digits\"[*>]*value=\"(?<calling_gt>\d+)\\"
| rex field=packet.proto.field.field.field[@showname]" "localValue:\s(?<action>[^\(]+)"
| rex field=_raw "<field name=\"gsm_map\msisdn\"[*>]*value=\"(?<msisdn>\d{10,15})\\"
| search "Component: invoke (1)" "localValue: anyTymeInterrogation (71)"
```

Figure 4.1: Definition of the correlation search associated with the ATI attack.

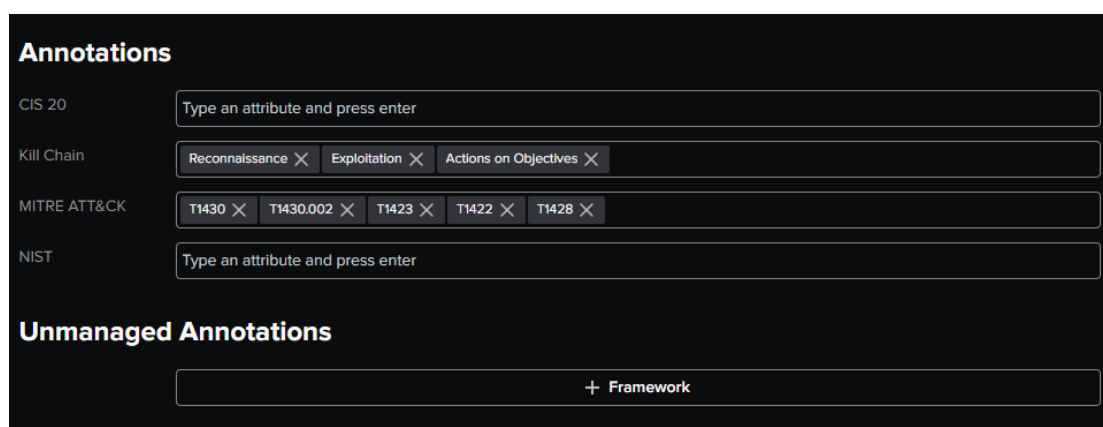
Regarding the RSI attack, the *name* of the correlation search has been defined in the same format as the ATI attack with a change specifying the appropriate attack type, resulting in Tndk - SS7 - RequestSubscriberInfo Attack Detection. The *application* and *application context* are, as the previous configuration, Enterprise Security. The *description* explains what the search is trying to detect, i.e. a GT abusing the SMS forwarding request to locate

a target. The mode is set to `manual` and the `search` is the one defined in listing 3.2.

## 4.2 Mapping

Splunk facilitates the association of a correlation search with various cybersecurity frameworks such as the Cyber Kill Chain [41] or the MITRE ATT&CK Mobile Matrix [42], but also supports the definition of custom frameworks. This allows the overall traceability of a detection, improving the knowledge related to the purpose of each of them. As the MITRE ATT&CK Matrix and Cyber Kill Chain are the most relatable frameworks to the nature of attacks over telecommunication networks, they will be used to perform the mapping.

Regarding the Kill Chain, the RSI and ATI attacks have been both associated with multiple phases: *Reconnaissance*, as the attacker has to retrieve the MSISDN of the target to begin the attack, but also *Exploitation*, as MAP messages are injected in the SS7 network taking advantage of its lack of security measures, and *Actions on Objectives*, as they both are centered around tracking the location of a subscriber. Regarding the MITRE ATT&CK Mobile Matrix, the correlation searches have been both associated with the techniques related to T1430 *Location Tracking* and T1430.002 *Impersonate SS7 Nodes*, T1423 *Network Service Scanning*, T1422 *System Network Configuration Discovery* and T1428 *Exploitation of Remote Services*. *Location Tracking* and *Impersonate SS7 Nodes* because the attacker is impersonating the MSC or the HLR to track the location of a subscriber in both attacks. *Network Service Scanning* and *System Network Configuration Discovery* as the attacker gains knowledge about the address location of the MSCs serving the target, which can be used in the future to carry out further attacks. Lastly, *Exploitation of Remote Services* due to the fact that both attacks are exploiting the SS7 trust mechanisms. Figure 4.2 showcases the aforementioned discussions mapped in the configuration of the correlation search.



The screenshot displays the 'Annotations' configuration page in Splunk. It features a dark-themed interface with a sidebar on the left and a main content area on the right. The sidebar lists four frameworks: 'CIS 20', 'Kill Chain', 'MITRE ATT&CK', and 'NIST'. The main content area shows the configuration for each framework. For 'CIS 20' and 'NIST', there is a text input field with the placeholder 'Type an attribute and press enter'. For 'Kill Chain', there are three buttons labeled 'Reconnaissance', 'Exploitation', and 'Actions on Objectives', each with a close icon (X). For 'MITRE ATT&CK', there are five buttons labeled 'T1430', 'T1430.002', 'T1423', 'T1422', and 'T1428', each with a close icon (X). Below the 'Annotations' section, there is a section titled 'Unmanaged Annotations' with a single button labeled '+ Framework'.

**Figure 4.2:** Configuration of the mapping with cybersecurity frameworks for both ATI and RSI attacks.

## 4.3 Timing

When referring to the timing of the correlation search, Splunk requires the specification of two categories of times. The first one is related to the time interval in which events should be taken into consideration and considers the source of the *timestamp*, either the ingestion time or the actual time the log has been created, and the *earliest* and *latest* time for the time range. The second time is related to the automation aspects and considers: the scheduling time, defined in a *cron schedule* format, the way the *scheduling* for the next search, the *schedule window* in which the search can be postponed or anticipated and its *schedule priority* among concurrent searches.

Both correlation searches have been configured in the same format, as they are both location tracking attacks and their characteristics are similar. It has been deemed appropriate to schedule the execution of the searches every six hours, resulting in four possible detection periods each day. Other than not causing an excessive load on the indexers that could delay the execution of other correlation searches, if the runs are scheduled at the appropriate time, it allows the relevant roles to assess the eventual presence of attacks in at least two rounds throughout the work day. Figure 4.3 showcases the configuration in common between the correlation searches:

**Time Range**

Timestamp ☒ Event time ☐ Index time  
Helps identify events that get included in the search results. Use index time only for raw event index correlation searches. [Timestamp documentation](#)

Earliest Time   
Set a time range of events to search. Type an earliest time using relative time modifiers.

Latest Time   
Type a latest time using relative time modifiers.

Cron Schedule   
Enter a cron-style schedule. For example `**/5 * * * *` (every 5 minutes) or `*0 21 * * *` (every day at 9 PM). Real-time searches use a default schedule of `**/5 * * * *`.

Scheduling ☒ Real-time ☐ Continuous  
Controls the way the scheduler computes the next execution time of a scheduled search. This controls the `realtime_schedule` setting. [Learn more](#)

Schedule Window   
Time interval during which to run the report.

Schedule Priority   
Setting to increase the priority of a report over other searches. Use with caution.

**Figure 4.3:** Time Range section of the correlation search for both the ATI and RSI attacks.

The *timestamp* has been set to Event Time as it is more efficient to make use of the ingestion time of Splunk rather than extracting the value from the field in the event. The *earliest time*

has been set to -6h and the *latest time* has been set to now to only consider the events of the past six hours. The *cron schedule* has been set to the 0 7/6 \* \* \*, instructing the search to run every six hours starting from 7:00. The *scheduling* is set to Real Time, as a continuous approach is more suitable for progressive flows of non packetized information. The *schedule window* is set to 30 minutes, allowing the search to be run within  $\pm 30$  minutes from its actual scheduled time. The *schedule priority* has been set to higher, which has a slightly elevated priority compared to the default one but lower than highest, allowing the search to not to be pushed back by others but at the same time avoid delaying the ones with an even higher priority.

## 4.4 Adaptive Response Actions

The remaining aspect to consider is related to the actions Splunk should take in case of detection. As a SIEM system is not a SOAR, its responses are only limited to passive notifications. The configuration needed for this step is related to: the definition of a *trigger* threshold and the *trigger amount* of times an action should be executed, the *window duration* in which the results can be throttled, i.e. the duplicates within the window are removed, the *throttle fields* having unique IDs that should be throttled, and the actions to be taken.

Even in this case, both correlation searches have been assigned the same configuration, apart from the naming conventions in the actions. Figure 4.4 showcases it:

The screenshot shows the 'Trigger Conditions' and 'Throttling' configuration sections in a dark-themed interface.

**Trigger Conditions:**

- Trigger alert when:** A dropdown menu showing 'Number of Results'.
- is greater than:** A dropdown menu showing 'is greater than' and a text input field containing '0'.
- Trigger:** Two radio buttons: 'Once' (selected) and 'For each result'.
- Notable response actions and risk response actions are always triggered for each result.

**Throttling:**

- Window duration:** A text input field containing '24' and a dropdown menu showing 'hour(s)'.
- Specify the time duration during which events that match the values specified in the "Fields to group by" might be ignored.
- Fields to group by:** A text input field containing 'msisdn' and 'calling\_gt', each with a close button (X).
- Type the fields to consider for matching events for throttling. [Learn more](#)

**Figure 4.4:** Defined Trigger Conditions.

Regarding the *trigger threshold*, an action should be taken when at least one event is detected, resulting in the greater than 0 condition. The *trigger amount* has been set to For each result to generate a notification for each detection instead of a general one. The *window duration* has been set to 24h and the *throttle fields* are the msisdn and the calling\_gt



fields which have been previously extracted in the searches. By setting up a throttling of one day, it allows the search to only notify attacks on new targets which avoids multiple notifications for the same combination of attacker and target.

Regarding the action to perform, out of the many possibilities provided by Splunk, the most applicable ones for the use case have been found in:

- *Email notifications* to a specific user or group of users.
- *Generation a Notable Event*, i.e. an event that can be used during audits.
- *Creation of a ticket* in the internal ticketing platform.

The most appropriate actions for the context of the correlation searches have been determined to be the email notification and the generation of a notable event. This is because the former allows to directly notify the appropriate roles focusing on the security of the telecommunication network which might not have direct access to Splunk. The latter allows instead the creation of particular events that are accessed during periodic audits, other than allowing the execution of further actions to gather more knowledge about the context. Regarding the creation of a ticket, even though its usage would improve the traceability of the actions that have been performed, additional internal knowledge related to the appropriate departments service accounts for the integration is needed, excluding it from the scope of the thesis.

#### 4.4.1 Structuring the Email Notification

Regarding the generation of an email, Splunk allows the usage of values associated with fields extracted from the search itself, other than attaching events in the email content. The email addresses of the appropriate roles managing the core network infrastructure have been specified in the destination field, *To* in the image. The *priority*, intended as importance in scheduling the sending operations, has been put to normal. The *Subject* has been set to `Splunk Alert:$name$`, in which the name of the correlation search replaces the `$name$` variable. The email *message*, mentions the the correlation search triggering the alert and contains information related to the IMSI of the target as well as the GT of the attacker. A PDF containing the event is also included in the email attachments, allowing the roles that do not have access to Splunk to take actions without having the right to access it. Figure 4.5 showcases the aforementioned email configuration:

Send email

To: lenor.dk, telenor.dk, lenor.dk, telenor.dk  
Comma separated list of email addresses. [Show CC and BCC](#)

Priority: Normal

Subject: Splunk Alert: \$name\$  
The email subject, recipients and message can include tokens that insert text based on the results of the search. [Learn more](#)

Message: an \$action\$ attack has been detected for the number \$msisdn\$ and generated by the GT \$calling\_gt\$.

Include:

- ☒ Link to Alert
- ☒ Link to Results
- ☐ Search String
- ☐ Inline
- ☐ Table
- ☐ Trigger Condition
- ☐ Attach CSV
- ☐ Trigger Time
- ☒ Attach PDF

Type: HTML and Plain Text Plain Text

Figure 4.5: Email configuration for for both RSI and ATI attacks.

#### 4.4.2 Structuring the Notable Event

Notable events are, as previously mentioned at the beginning of this chapter, a particular type of events that is used as the starting point to perform further investigation. When the threshold condition is met, a new entry with its own *title* and *description* is created in a particular index in Splunk that is used to store events worthy of a more detailed review. A classification in terms of *security domain* and *severity* is also provided, along with an *owner* that should further investigate the issue and a *status* representing the progress on it.

Once a notable event is created, it is also possible to run further *drilldown searches* and update *drilldown dashboards*, which results are attached to the notable event in order to further provide context that could aid in the investigation process. It is also possible to attach the information contained in certain fields related to identify fields containing names or identities, the assets involved in the event, files, and Uniform Resource Locator (URL) related to the detection. These are respectively mapped in the notable event as *identity extraction*, *asset extraction*, *file extraction*, *URL extraction*. An association with an ongoing *investigation profile* can also be provided, which is used to group relevant notable events related to an

investigation under the same umbrella is also present. Lastly, *next steps* and *recommended actions* can also be provided in order to support the operations of known problems by specifying internal processes or procedures to be followed.

Even in this case, as both correlation searches are performing detections on the related scope, they have been configured in the same format with minor naming alterations. Figure 4.6 showcases the configuration of the notable event for both attacks.

The *title* has been set to `$action$ attack on $msisdn$`, showing the name of the MAP procedures being invoked and the MSISDN of the target. The *description* has been set to contain the time in which the attack occurred, specifying the GT of the attacker, the IMSI of the subscriber, and the type of attack performed. The *security domain* has been set to `Network`, as the detection has been performed on data originating from the network infrastructure. The *severity* has been set to `low`, as the detected attacks do not actively interfere with the activities of the target like an SMS interception would. A further increase of the priority would also imply greater attention needs, which might remove resources from the investigations of problems related to more important attacks. The *default owner* and *default status* have been left as system default and no *drilldown searches* or *drilldown dashboards* have been specified. The *investigation profile* has also been left to the default value, as no processes have been defined for these investigations. The *identity extraction* has been set to include the `msisdn`, `calling_gt` fields extracted from the searches and the *asset*, *file* and *URL extractions* have been left to their default values as they are not relevant from the scope of the correlation search. If such searches would be implemented based on the logs of real data, the identity of the SP in the core infrastructure forwarding the log could also be added in *asset extraction*. *Next steps* and *recommended actions* have also been left untouched.

+ Add New Response Action

Notable

Title

\$action\$ attack on \$mssisdn\$

Notable events created by this search will have this title. Supports variable substitution.

Description

an \$action\$ attack has been detected at \$\_time\$ for the number \$mssisdn\$ and generated by the GT \$calling\_gt\$.

Notable events created by this search will have this description. Supports variable substitution.

Security Domain

Network

Severity

Low

Used to calculate urgency for notable events. [Learn more](#)

Default Owner

(leave as system default)

Default Status

(leave as system default)

Drill-down Searches

+ Add Drill-down Search

Drill-down Dashboards

+ Add Drill-down Dashboard

Investigation Profiles

Select a profile.

Identity Extraction

src\_user user src\_user\_id src\_user\_role user\_id user\_role vendor\_account

Asset Extraction

src dest dvc orig\_host

File Extraction

Type a field name.

URL Extraction

Type a field name.

Next Steps

Insert Action

Describe the next steps and adaptive response actions to address this threat. Add the URL using the following syntax: `[[actionNameOfAction]]`. [Learn more](#)

Recommended Actions

All

Search...

Recommended

Search...

Identifying Recommended Adaptive Responses will highlight those actions for the analyst when looking at the list of response actions available, making it easier to find them among the longer list of available actions.

Figure 4.6: Configuration of the generation of a notable event for both ATI and RSI attacks.

## Chapter 5

# Conclusion

This thesis had the goal of improving the security of telecommunication networks by making use of a SIEM system to correlate the messages exchanged in the core infrastructure and detect the presence of attacks leveraging the MAP protocol of the SS7 suite, as stated in the final problem statement:

*How to define searches in Splunk to detect the presence of SS7 attacks over MAP in the logs produced by the core network infrastructure, while maintaining acceptable detection levels with a limited amount of input sources?*

As discussed in section 3.1, the usage of network data generated by the real SPs of the network infrastructure was not possible, both in terms of anonymized information and publicly available datasets. The adopted approach therefore consisted of making use of the JSS7 simulator to generate network logs mimicking the behavior of a real telecommunication network. Three isolated scenarios consisting of location tracking and SMS interception attacks were created, but also a complex simulation in which they were blended with normal interactions between subscribers. Such simulations have then been ingested in Splunk in XML like structure, the *.pdml* file format, which allowed the ingestion of events containing the entire packet information with a tradeoff in terms of requiring manual extraction of the necessary values from each field using regular expressions.

To facilitate the identification of the attack indicators, as discussed in section 3.3, Wireshark was used to explore the content of specific SS7 packets in a more human intelligible format than a XML structure. The resulting findings have then been translated into searches defined in SPL which allowed the detection of location tracking attacks leveraging the MAP ATI messages or the combination of the MAP RSI and PSI requests. The defined searches present in listing 3.1 and listing 3.2 have first been tested on the isolated attack datasets and later in the complex simulation. This ensured that the defined searches were actually able to detect the presence of attacks, and later identify them in a realistic scenario.

The ATI search was able to detect all of the 248 MAP ATI messages in the complex simulation without the presence of false negatives. Such results were possible to achieve for two different reasons: GSMA suggests the network operators to consider each ATI call as malicious especially if originating from outside their PLMN, as described in subsection 3.3.1, but also because the ATI attack pattern only consists of isolated ATI invocations, as shown in Figure 2.6. The RSI attack was able to only perform the detection in the isolated setting due to the absence of actual attacks of such type in the realistic scenario, as discussed in subsubsection 3.3.2. This has also been manually verified by analyzing the network logs using Wireshark.

The resulting searches have then been incorporated in correlation searches that periodically run four times per day, as described in chapter 4. This enables the appropriate roles to initiate an investigation on the eventual detections generated each day. The notification methods after the identification of an attack have been defined to include the generation of an email to the appropriate roles as well as the creation of a notable event that will be used during the actual investigation of such threats.

It is worth mentioning that the RSI search is only able to detect the presence of attacks when the RSI message is followed within one second by a RSI invocation. If an attacker already has knowledge of the MSC serving the target, which could have been found by performing the RSI attack at different times or by acquiring such information with other tactics. By directly sending a PSI message to the MSC serving the target, the attacker would be able to bypass the detection.

For these reasons, future work will focus on, other than gaining access to real network data, the extension of the definition of the correlation searches for the SMS interception attacks as well as other attacks on the SS7 stack described in subsection 2.4.2. Another possible path that could be investigated could be found in applying Machine Learning to achieve a dynamic detection of attack patterns in the signaling flows whose steps are performed at different times.

Regarding the applicability of the thesis work to a real data, supposing that the JSS7 simulator mimics the SS7 protocol suite in a 1:1 mode, major changes in the behavior of the defined searches should not be needed. This is because the attack patterns to be identified for both of them have been found through an analysis of the of the MAP and SCCP protocol headers. However, changes in the representation format are to be expected, as a live flow of packets might not make use of an XML like structure to represent the information. A different data format might also allow Splunk to perform automatic field extraction, thus stripping the searches of the regular expressions needed to perform such action.

## 5.1 Input Sources and Flow of information

The final problem statement also mentioned the relation with a limitation on the number of input sources and acceptable detection levels. The aspects related to licensing and storage costs, as described in subsection 2.1.1, is a factor that needs to be taken into consideration when expanding the operational capacities of a SIEM system.

If considered that the number of customers for Telenor Denmark in 2025 was around 1.7 million [43], along with a difficult to estimate number of roaming devices, the amount in terms of signaling messages to be stored and analyzed must be considered to determine the applicability of the thesis work. In order to detect ATI and RSI attacks, it is necessary to obtain the network logs from each MSC and HLR present in the infrastructure. If more complex attacks should also be detected, it might be necessary to expand the logging activities to all the SPs.

Lemaire et al. [44] estimates the collection of such signaling data at a level of detail up to BS level related to 27 million subscribers operating in the French metropolitan areas to be in the order of 12 Terabytes per day. If this estimation is translated to a smaller national network with on average 2 million subscribers, the data volume related to the signaling message could be in the order of 4.5 Terabytes per day, which is 90 Terabytes per month and roughly 1.65 Petabytes per year. This amount of information, other than requiring infrastructural expenses for its storage and security, implies a further consideration of the licensing costs of ingesting such data volumes in Splunk.

In the case of Telenor, the current subscription plan for Splunk is related to the amount of data volumes ingested per month. Such an increase in the data inputs would therefore cause an exponential rise in its costs, reducing the overall feasibility of the thesis project. Out of the plans offered by Splunk, as described in subsection 2.1.1, the most optimal arrangement in relation with the scope of the project would therefore be based on the number of entities forwarding their logs. This is because its license costs can be modulated based on the SPs deemed important for logging an attack, thus not having limitations on both ingestion volumes or workloads necessary to analyze them.

Infrastructural changes of such importance would however require years to be completely implemented, as it is necessary to undergo various approval processes, acquire the necessary infrastructure, and allocate the necessary resources to successfully implement the detections.

## 5.2 Acknowledgments

I would like to express my gratitude to Telenor Denmark for allowing me to conclude my academic career with the opportunity of learning and applying my knowledge on the topic of cybersecurity over telecommunication networks. I would also like to thank my company supervisor Niels and my university counterpart Tatiana for their support and useful guidance throughout the development of this thesis.



# Bibliography

- [1] Lopa J Vora. “Evolution of mobile generation technology: 1G to 5G and review of upcoming wireless technology 5G”. In: *International journal of modern trends in engineering and research* 2.10 (2015), pp. 281–290.
- [2] Mohsen Attaran. “The impact of 5G on the evolution of intelligent automation and industry digitization”. In: *Journal of ambient intelligence and humanized computing* 14.5 (2023), pp. 5977–5993.
- [3] Peng Liu, Thomas F. LaPorta, and Kameswari Kotapati. “Cellular Network Security”. In: *Network and System Security*. Ed. by Elsevier Inc. Elsevier Inc., 2013, pp. 319–351. ISBN: 9780124166899. DOI: 10.1016/B978-0-12-416689-9.00011-3.
- [4] 3GPP. *5G System Overview*. 2022. URL: <https://www.3gpp.org/technologies/5g-system-overview>.
- [5] Jani Suomalainen et al. “Securing Public Safety Communications on Commercial and Tactical 5G Networks: A Survey and Future Research Directions”. In: *IEEE Open Journal of the Communications Society* 2 (2021), pp. 1590–1615. DOI: 10.1109/OJCOMS.2021.3093529.
- [6] Spyridon Samonas and David Coss. “The CIA strikes back: Redefining confidentiality, integrity and availability in security.” In: *Journal of Information System Security* 10.3 (2014).
- [7] Augusto Barros, Anton Chuvakin, and Anna Belak. *Applying Network-Centric Approaches for Threat Detection and Response*. Tech. rep. Gartner, 2019. URL: <https://www.gartner.com/en/documents/3904768>.
- [8] Karen Ann Kent and Murugiah Souppaya. “Guide to Computer Security Log Management:.” In: (2006).
- [9] David R Miller. *Security information and event management (SIEM) implementation*. McGraw-Hill Higher Education, 2011.
- [10] Gustavo González-Granadillo, Susana González-Zarzosa, and Rodrigo Diaz. “Security information and event management (SIEM): analysis, trends, and usage in critical infrastructures”. In: *Sensors* 21.14 (2021), p. 4759.

- [11] Splunk Inc. *Splunk: The Data Platform for the Hybrid World*. Accessed: 2025-02-25. 2025. URL: <https://www.splunk.com/>.
- [12] David Carasso. *Exploring Splunk*. Cito Research New York, 2012.
- [13] Splunk. *Splunk Documentation*. 2025. URL: <https://docs.splunk.com/Documentation/Splunk>.
- [14] Splunk Inc. *Splunk Pricing FAQ*. [https://www.splunk.com/en\\_us/products/pricing/faqs.html](https://www.splunk.com/en_us/products/pricing/faqs.html). Splunk Website. 2025.
- [15] *The Cyber Threat Against Denmark 2024*. Tech. rep. Centre for Cyber Security, 2024. URL: <https://www.cfcs.dk/globalassets/cfcs/dokumenter/trusselsvurderinger/en/cfcs---the-cyber-threat-against-denmark-2024.pdf>.
- [16] European Union Agency for Cybersecurity (ENISA). *ENISA Threat Landscape 2024*. 2024. URL: <https://www.teledcis.dk/wp-content/uploads/2024/11/public-ENISA-Threat-Landscape-2024.pdf>.
- [17] *FS.57 Mobile Threat Intelligence Framework (MoTIF) Principles*. Tech. rep. GSMA, 2024. URL: <https://www.gsma.com/solutions-and-impact/technologies/security/wp-content/uploads/2024/04/FS.57-MoTIF-Principles-v1.0.pdf>.
- [18] European Union Agency for Cybersecurity (ENISA). *Signalling Security in Telecom SS7/Diameter/5G*. Tech. rep. ENISA, 2018. URL: <https://www.enisa.europa.eu/sites/default/files/publications/Interconnect%20Security%20SS7-Diameter.pdf>.
- [19] GSMA. *Mobile Telecommunications Security Landscape 2025*. 2025. URL: <https://www.gsma.com/solutions-and-impact/technologies/security/gsma-mobile-telecom%20munications-security-landscape-2025/>.
- [20] Shigeo Matsuzawa et al. “Architecture of cell switch router and prototype system implementation”. In: *IEICE transactions on communications* 80.8 (1997), pp. 1227–1238.
- [21] James F. Kurose and Keith W. Ross. *Computer Networking: A Top-Down Approach*. 8th. Pearson, 2020. ISBN: 978-0135928615.
- [22] Gunnar Heine and Matt Horrer. *GSM networks: protocols, terminology, and implementation*. Artech House, Inc., 1999.
- [23] Giuseppe Cattaneo, Giancarlo De Maio, and Umberto Ferraro Petrillo. “Security Issues and Attacks on the GSM Standard: a Review.” In: *J. Univers. Comput. Sci.* 19.16 (2013), pp. 2437–2452.
- [24] Kaleem Ullah et al. “SS7 vulnerabilities—a survey and implementation of machine learning vs rule based filtering for detection of SS7 network attacks”. In: *IEEE Communications Surveys & Tutorials* 22.2 (2020), pp. 1337–1371.

- [25] Lee Dryburgh and Jeff Hewett. *Signaling System No. 7 (SS7/C7): Protocol, Architecture, and Services*. Cisco Press, 2004. ISBN: 9781587050404. URL: <https://www.ciscopress.com/store/signaling-system-no.-7-ss7-c7-protocol-architecture-9781587050404>.
- [26] Håkan Ventura. *Diameter: Next generations AAA protocol*. 2001.
- [27] Hwankuk Kim. “5G core network security issues and attack classification from network protocol perspective.” In: *J. Internet Serv. Inf. Secur.* 10.2 (2020), pp. 1–15.
- [28] Telenor Norge. *2G-nettet hos Telenor*. 2024. URL: <https://www.telenor.no/dekning/2g/>.
- [29] Teleguiden. *2G-netværk lukkes i Danmark*. 2024. URL: <https://teleguiden.dk/2g-netvaerk-lukkes-danmark/>.
- [30] Martin Sauter. *From GSM to LTE-Advanced Pro and 5G: An Introduction to Mobile Networks and Mobile Broadband*. 4th ed. John Wiley & Sons, 2021. ISBN: 978-1-119-71467-5.
- [31] GSMA. *FS.07: SS7 and SIGTRAN Network Security (v4.0)*. 2025. URL: [https://www.gsma.com/solutions-and-impact/technologies/security/gsma\\_resources/fs-07-ss7-and-sigtran-network-security-v4-0/](https://www.gsma.com/solutions-and-impact/technologies/security/gsma_resources/fs-07-ss7-and-sigtran-network-security-v4-0/).
- [32] European Union. *Regulation (EU) 2016/679 of the European Parliament and of the Council*. 2016. URL: <https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng>.
- [33] Abdul Majeed and Sungchang Lee. “Anonymization techniques for privacy preserving data publishing: A comprehensive survey”. In: *IEEE access* 9 (2020), pp. 8512–8545.
- [34] Polarking. *SS7 Attack Simulator based on RestComm’s jss7*. <https://github.com/polarking/jss7-attack-simulator>. 2025. URL: <https://github.com/polarking/jss7-attack-simulator>.
- [35] Loay Abdelrazek. *SigPloit - Telecom Signaling Exploitation Framework*. <https://github.com/SigPloiter/SigPloit>. 2018.
- [36] RestComm. *JSS7*. 2025. URL: <https://github.com/RestComm/jss7>.
- [37] Yuejun Guo et al. “An Empirical Study of Deep Learning-Based SS7 Attack Detection”. In: *Information* 14.9 (2023), p. 509.
- [38] Yuejun Guo. “Simulated dataset for SS7 attack detection”. In: (July 2023). DOI: 10.6084/m9.figshare.23666397.v1. URL: [https://figshare.com/articles/dataset/Simulated\\_dataset\\_for\\_SS7\\_attack\\_detection/23666397](https://figshare.com/articles/dataset/Simulated_dataset_for_SS7_attack_detection/23666397).
- [39] Wireshark Foundation. *Wireshark: Network Protocol Analyzer*. <https://www.wireshark.org/>. 2025.

- [40] GSM Association. *FS.11 SS7 Interconnect Security Monitoring and Firewall Guidelines*. 2019. URL: [https://www.gsma.com/solutions-and-impact/technologies/security/gsma\\_resources/fs-11-ss7-interconnect-security-monitoring-and-firewall-guidelines-v6-0/](https://www.gsma.com/solutions-and-impact/technologies/security/gsma_resources/fs-11-ss7-interconnect-security-monitoring-and-firewall-guidelines-v6-0/).
- [41] Eric M Hutchins, Michael J Cloppert, Rohan M Amin, et al. "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains". In: *Leading Issues in Information Warfare & Security Research* 1.1 (2011), p. 80.
- [42] MITRE. *MITRE ATT&CK® Mobile Matrix v17*. 2025. URL: <https://attack.mitre.org/versions/v17/matrices/mobile/>.
- [43] Telenor A/S. *Telenor fortsætter fremgangen: Stærkt første kvartal med omsætningsvækst og øget kundetilgang*. 2025. URL: <https://press.telenor.dk/pressreleases/telenor-fortsaetter-fremgangen-staerkt-foerste-kvartal-med-omsaetningsvaekst-og-oeget-kundetilgang-3384448>.
- [44] Pierre Lemaire et al. "Early detection of critical urban events using mobile phone network data". In: *Plos one* 19.8 (2024), e0309093.

## Appendix A

# Searches for Displaying Purposes

```
index = main host=SS7TestData source="SS7 Capture.pdml"
| rex field=_raw "<field name=\"sccp\.calling\.digits\"[^>]*
  value=\"(?<calling_gt>\d+)\ \""
| rex field=_raw "showname=\"localValue:\s*(?<action>.*?)\s?\\(\\
  d+\\)\ \""
| rex field=_raw "<field name=\"gsm_map\.msisdn\"[^>]*value
  =\"(?<msisdn>\d{10,15})\ \""
| search "Component: invoke (1)" "localValue:
  anyTimeInterrogation (71)"
| table _time, calling_gt, action, msisdn
```

**Listing A.1:** Splunk search based on listing listing 3.1 that displays a table for the ATI attack.

```
index = main host=SS7TestData
| rex field=_raw "name=\"sccp\.calling\.digits\".*?value=\"(?<
  calling_gt>\d+)\ \""
| rex field=_raw "showname=\"localValue:\s*(?<action>.*?)\s?\\(\\
  d+\\)\ \""
| rex field=_raw "<field name=\"gsm_map\.(?:sm\\.)?msisdn\"[^>]*
  value=\"(?<msisdn>\d{10,15})\ \""
| transaction calling_gt maxspan=1s
| search "localValue: sendRoutingInfoForSM (45)" AND "
  localValue: provideSubscriberInfo (70)" AND "Component:
  invoke (1)"
| table _time, calling_gt, action, msisdn
```

**Listing A.2:** Splunk search based on listing listing 3.2 that displays a table for the RSI attack.