CHALMERS | UNIVERSITY OF GOTHENBURG


# A First Security Analysis of a Secure Intermodal Goods Transport System

JONAS ANGHOLT
MIKAEL WACKERBERG
TOMAS OLOVSSON
ERLAND JONSSON

Department of Computer Science and Engineering
*Division of Networks and Systems*
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Göteborg, Sweden, 2013-09-04
Technical Report No. 2013:05
ISSN 1652-926X

# Foreword

The overall project *Secure Intermodal Transport System* (SITS) is addressing the necessity of improving security within the transport industry as outlined in the pre-study report *"The Benefits of a Standardized Communication Interface for Secure Intermodal Transports"*. Here, various threats, such as smuggling, terrorism theft and other incidents, are described. The report also discusses how improved communication facilities could potentially mitigate the consequences of such threats. The SITS project is also partly based upon the previous project *"Secure and Efficient Transports through Ports"*.

This report specifically makes an analysis of the security issues of two communication entities: the back-office communication and the RFID-based system for automatic check-in at terminals. The analysis is preliminary as the main SITS project has suffered from substantial delays, and in the absence of the final design it was not possible to draw definitive conclusions. However, it is believed that our results reflect the security of the system to a satisfactory degree. The work has been carried out at the department of Computer Science and Engineering, mainly by two Master's thesis students.

# Abstract

The goods transport business involves a lot of money and is a big part of the infrastructure of any European country. There are often many different actors involved in each transport and the communication network is rather complex due to the point-to-point communication structure. It is easy to understand why there is a high demand for increased simplicity and effectiveness. With this in mind, the e-Freight project which is based on PEPPOL has moved towards a standardized solution by developing a communication system based on access points (APs). These APs act as the interface to the system and makes it easy to establish communication between any two connected actors.

With PEPPOL and e-Freight as a foundation, VOLVO leads the SITS project in close cooperation with Stena Line and DSV. The goal is to develop a harmonized communication framework that promotes increased sharing of information between actors and enable new applications to increase effectiveness and security in the chain of transportation. This leads to simplified accessibility for actors to a set of services by being connected to an AP. At the same time service providers benefit from being able to easily set up cloud services available for all actors. In addition to the back-office communication between APs, external devices such as cellphones, in-vehicle computers and check-in terminals can communicate directly with each other over short distances. This type of communication is only partially specified and a mutual standard is yet to be decided upon.

In this report we have analyzed the SITS project from an IT-security perspective. The back-end system derived from e-Freight is looked into and communication links, access points, protocols, certificate handling etc., are examined. Another concern in the SITS project is the short-range communication between trucks and terminals. Since RFID is a highly potential candidate for use in this area, we have studied the technology by categorizing a typical RFID system into three distinct layers and researched important security threats with the classic CIA approach. Based on the security issues found, countermeasures such as encryption, authentication and protection against man-in-the-middle attacks are reviewed.

Keywords: PEPPOL, BusDox, e-Freight, SITS, RFID, security, confidentiality, integrity, availability, encryption, authentication, man-in-the-middle, certificate

# Contents

# Abbreviations

| | |
|---|---|
| **AP** | Access Point |
| **BMS** | Business Message Standard |
| **BusDox** | Business Document Exchange Network |
| **CA** | Certificate Authority |
| **CIA** | Confidentiality, Integrity and Availability |
| **CIAAA** | Confidentiality, Integrity, Availability, Authenticity and Accountability |
| **CRL** | Certificate Revocation List |
| **DDoS** | Distributed Denial of Service |
| **DNS** | Dynamic Name System |
| **DNSSEC** | Domain Name System Security Extensions |
| **DPA** | Differential Power Analysis |
| **EMA** | Electromagnetic Analysis |
| **EPC** | Electronic Product Code |
| **ETSI** | European Telecommunications Standards Institute |
| **GDP** | Gross Domestic Product |
| **HMAC** | Hash-based Message Authentication Code |
| **HTTP** | Hypertext Transfer Protocol |
| **IDS** | Intrusion Detection System |
| **IETF** | Internet Engineering Task Force |
| **LIME** | Lightweight Message Exchange |
| **MAC** | Message Authentication Code |
| **MITM** | Man in the Middle |
| **NFC** | Near Field Communication |
| **OASIS** | Organization for the Advancement of Structured Information Standards |
| **ONS** | Object Name Service |
| **PEPPOL** | Pan-European Public Procurement On-Line |
| **PKI** | Public-key Infrastructure |
| **RFID** | Radio-Frequency Identification |
| **ROM** | Read Only Memory |
| **SML** | Service Metadata Locator |
| **SMP** | Service Metadata Publisher |
| **SAML** | Security Assertion Markup Language |
| **SITS** | Secure Intermodal Transport System |
| **SPA** | Simple Power Analysis |
| **SQL** | Structured Query Language |
| **SSL** | Secure Socket Layer |
| **START** | Secure Trusted Asynchronous Reliable Transport |
| **TLS** | Transport Layer Security |
| **TVRA** | Threat, Vulnerability and Risk Analysis |
| **W3C** | World Wide Web Consortium |
| **Wi-Fi** | Wireless Fidelity |
| **XML** | Extensible Markup Language |

# 1 Introduction

## 1.1 Background

The importance of global trade is constantly increasing and is for many countries heavily contributing to the gross domestic product (GDP). This is partly due to the establishment of the World Trade Organization in 1995 which goal is to promote free trade and at the same time create a globally regulated trade structure. Along with the increased possibilities of trade, the regulations concerning transportation of goods have growing significantly more complex. When shipping across borders there are typically 35 documents that need to be exchanged between 25 parties. This is the result of regulations being developed separately in different countries and for different modes of transportation. That variety represents barriers for the industry and causes companies to face a complex set of excessive and redundant reporting requirements. This makes it necessary to develop and maintain interfaces with multiple national systems. This causes great costs for all parties both in terms of money but also when it comes to effectiveness and reliability [24].

There have been other identified problems beside the complexity of regulations. One example is the fact that a lot of information involved in the transportation of goods is handled manually via paper documents. This can at times be very time consuming, such as when a truck driver arrives at a terminal and has to leave the vehicle to bring all the required documentation to the terminal operator. Another type of problems are antagonistic threats such as smuggling, vandalism and theft, both around terminals and along roads. There is also a problem concerning the management of transports, especially when it comes to goods with restriction. It can be hard to plan and effectively monitor and control dangerous goods transports, something that becomes evident when an accident occurs and information needs to be available and authorities needs to be alerted as fast as possible [9].

To address these problems, Volvo Technology has started a project called SITS (Secure Intermodal Transport System). The aim of the project is to propose a communication framework that enables electronic sharing of information between actors in the goods transport industry. Volvo will construct two specific business cases: automated check-in at terminals and anomaly detection mainly looking for accidents and illegal activity [16].

## 1.2 Aim

The aim of this project is to perform a security analysis of the SITS project. We have established a description of the target and using this we have identified the security concerns and potential countermeasures to be implemented in the future. We have tried to answer the following questions in our report:

Which are the assets in the system that needs to be protected?

Which are the threats against the system? Which vulnerabilities can a potential hacker exploit?

Which security countermeasures should be implemented?

## 1.3 Scope and limitation

The scope of this report is limited to the IT security of the SITS project and thus we do not consider how to physically secure assets such as the goods being transported on a truck. Out of the two business cases defined in SITS we have only focused on automatic check-in and do not consider the anomaly detection system. However, communicating with the anomaly system should be similar to communicating with any other business partner so that part of the analysis should still be relevant to this business case.

The security analysis of the proposed short range communication solutions is limited to the ones involving RFID. Furthermore, this report does not consider the security of the internal systems of the involved parties, which is also the limitation of the SITS project itself.

## 1.4 Assumptions

As the SITS project is still in an early part of its development many of the technical specifications are far from decided. However, since the aim is to adapt the communication infrastructure from the PEPPOL project, we assume that the final specifications for the SITS project will not differ significantly from PEPPOL.

# 2 Method

In order to perform a systematic risk-analysis of SITS we have chosen to use an approach called CORAS (which, even though it is written in all capital letters, is actually not an acronym). CORAS consists of three different parts: a language, a tool and a method.

The language is diagrammatic and consists of graphical symbols and relations between them to create easily understood diagrams. The diagrams are created and used during brainstorming sessions to document the discussion.

The tool is an editor that is used to create CORAS diagrams and can be downloaded for free from the CORAS project homepage.

The CORAS method is an asset-driven method for risk-analysis which is composed of eight steps. In risk-analysis an asset is something of value that needs to be protected. In a method that is asset-driven these assets are identified early on in the process and the direction of the analysis is then driven by them.

The CORAS approach was chosen as the method for this analysis largely because of the excellent documentation of it in the form of the textbook "Model Driven Risk Analysis". The book describes not only *what* is to be done in the risk analysis but also gives instructions as to *how* the different parts of the analysis can be performed, which we felt would be very valuable since neither of us has performed a risk analysis prior to this. The examples provided in the book are also more relevant since they almost exclusively deal with the domain of IT security as opposed to safety and mechanical systems which is the focus of much of the literature on risk analysis.

## 2.1 Terminology and Symbols

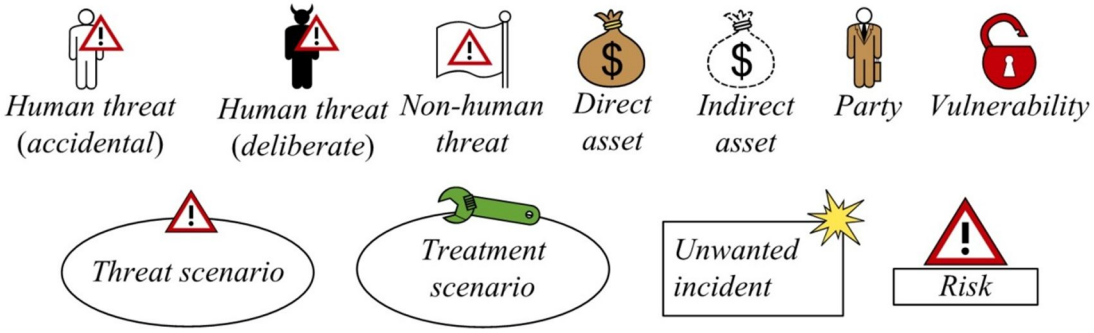


**Figure 2.1: CORAS symbols**

The first three symbols in the figure represent different kinds of threats. In the CORAS language a threat is defined as "a potential cause of an unwanted incident". Threats are divided into three different types: Human threat (accidental), Human threat (deliberate) and Non-human threat.

A *deliberate human threat* is a human being with a malicious intent. A good example of this kind of threat (and of the concept of a threat in general) is a hacker.

A human threat does not however need to someone with a malicious intent. *Accidental human threats* are, as the name implies, also human threats but refer more specifically to those who have no intent of causing harm, e.g. a system user who, due to lack of competence, makes an error that causes an unwanted incident.

Threats to a system are not necessarily human, thus the third type of threat that we consider is the *non-human threat*. This threat covers all of the ones that are not human in nature such as computer viruses or threats that are out of our control such as natural disasters.

CORAS defines an *asset* as "something to which a party assigns value and hence for which the party requires protection". The definition is purposefully general in order to include everything that an involved party may wish to protect. Assets in CORAS are divided into two types: direct assets and indirect assets. *Direct assets* are those that can be harmed directly by an unwanted incident. A direct asset may for example be a physical object, such as a computer, or it may be the data that is stored on the computer. It can also be different properties of an entity, e.g. the availability of a server may be one asset while the integrity of the same server is another asset. An indirect asset on the other hand is something that may only be harmed via other assets. A typical example of this would be a company's reputation which could for example be harmed by their customers' sensitive information being stolen by a hacker.

A *party* is someone who has an interest in the target of the analysis, such as an organization, a person and some other stakeholder. In most cases there is only one party, which is the same as the one who commissions the risk-analysis. In some cases though (such as ours) there are several different parties to consider during the risk analysis who may have defined different assets or have the same assets but they assign differing values to them.

A *vulnerability* is a weakness in the system which a threat may exploit in order to cause harm to an asset. An example of this would be a server which does not have an anti-virus program installed which enables a virus to infiltrate the server.

While a threat is *what* may cause harm to an asset, a *threat scenario* is the description of *how* the threat may harm it. This is modeled as series of events that are initiated by a threat and ultimately leads to an unwanted incident. Each event contains a brief description of the event with the level of detail depending on the scope of the analysis.

A *treatment scenario* is the implementation of a countermeasure that is intended to reduce risk. The most common and perhaps most obvious way of reducing risk is to aim the countermeasure at a vulnerability but a treatment scenario may in fact be aimed at any element in a threat diagram. We could for example address the vulnerability "lack of virus protection" by installing anti-virus software thus reducing the likelihood of being infected with a virus or we may have identified "untrained employees" as a threat and we may deal with that by implementing the treatment scenario "train employees".

An *unwanted incident* is defined as "an event that harms or reduces the value of an asset". This could be the unwanted incident "server goes down" which would harm the asset "availability of server". There may also be cases where an unwanted incident causes harm to multiple assets and where an unwanted incident causes another unwanted incident.

Loosely, a *risk* is the likelihood of an unwanted incident occurring multiplied by the consequence of it. Since the consequence of an unwanted incident is dependent upon the asset which it harms there is one risk associated with every combination of unwanted incident and asset. Each risk also has an associated risk level which is derived from the likelihood and consequence, this risk level is used in the analysis to determine which risks need to be addressed with countermeasures.

## 2.2 The CORAS Method

In this section we will briefly describe the steps of the CORAS method. It should be noted that "customer" here refers to the organization that has ordered the risk-analysis and "analysts" refers to the team who is contracted to perform the analysis.

**Step 1 - Preparations for the analysis**
The first step of the CORAS method is to meet with the customer to do some preparatory work before the actual risk analysis begins. The customer provides the analysts with background information and documentation that is relevant to the analysis and a rough outline of the scope of the analysis is suggested.

**Step 2 - Customer Presentation of the Target**
In the second step a meeting is held where the customers give a presentation of the target of the analysis and what the goals of the analysis are. In order to minimize the chance of misunderstandings between the analysts and the customer a presentation of the CORAS approach is also given and the terminology that is used in the risk analysis is presented.

**Step 3 - Refining the Target Description Using Asset Diagrams**
The third step is conducted to ensure that the analysts and the customer have a common understanding of the target of the analysis. The analysis team will present the target of analysis based on their understanding of the documentation given to them in step 1 and the presentation given by the customer in step 2. The customer can then correct or clarify certain aspects if needed. If not, then the meeting continues by identifying the important parts of the system that needs to be protected, i.e. the assets. This is documented with an asset diagram using the CORAS tool. After this, a high-level analysis is performed to identify the major threats and vulnerabilities. This overview then helps in further defining the scope of the analysis.

**Step 4 - Approval of the Target Description**
The fourth step of the CORAS method is the last step before the actual risk analysis starts. In this step the analysts and the customer agree on a final description of the target, including the scope and the assumptions that are made. This step also includes deciding what scales are to be used when estimating risks and consequences.

**Step 5 - Risk Identification Using Threat Diagrams**
The fifth step is conducted as a workshop where the analysts with the help of people from different backgrounds try to identify the risks that are relevant to the target. This includes identifying the threats to the system, the unwanted incidents that could harm the assets, the threat scenarios that may lead to these unwanted incidents and the vulnerabilities that could be exploited by a threat. The results are documented during the workshop in a threat diagram.

**Step 6 - Risk Estimation Using Threat Diagrams**
The aim of the sixth step is to analyze the results from the previous step to estimate the likelihood of the unwanted incidents to occur and the consequences if they do. By combining the likelihood and consequence we get the risk level for each unwanted incident.

**Step 7 - Risk Evaluation Using Risk Diagrams**
During the seventh step we evaluate the risks that we estimated in the previous step to decide which ones are acceptable and which ones we need to investigate treatment methods for.

**Step 8 - Risk Treatment Using Treatment Diagrams**
At the eight step the risks that were found to be unacceptable in step seven are analyzed and treatments for them are identified. This can mean either reducing the likelihood of the incident occurring or the consequences if it does occur (or both). Furthermore, the treatments that are found are also analyzed from a cost-benefit perspective to ensure that the costs of treatments are cost-effective.

# 3 The SITS Framework

The SITS project is a proposition for a communication framework to be used in the exchange of information between different parties in the transport industry. Currently the communication between parties is peer-to-peer where each company has their own interface. The number of communication links can therefore be more than wanted for a large or medium sized company. The SITS framework addresses this problem by suggesting a unified system relaying all communication. The project is headed by Volvo Technology in cooperation with Stena Line, DSV, Chalmers, Saab, MSB (Swedish Civil Contingencies Agency) and Wackfelts Transport. In order to formulate the requirements of the framework, two conceptual solutions have been developed in parallel with the framework. The first is a system for automatically handling the check-in at a terminal. The second one is a system developed by Saab for identifying anomalies, such as detecting an accident or a transport which matches the custom offices risk profile, and informing the concerned authorities. However, this report only focuses on the automatic check-in scenario.

## 3.1 Related projects

In this section we give a brief outline of projects that relate to SITS in some way.

### 3.1.1 Port pilot

Port pilot was a pilot project led by Volvo Technology on behalf of Lindholmen Science Park Security Arena. In the project a truck was outfitted with the technology to transmit information automatically at the check-in at transport terminals instead of the driver leaving the truck and enter the terminal building to authorize herself and hand over the necessary documentation. This meant that the time the truck was left unattended was eliminated during the check-in process. The truck was also equipped with a scanner that could identify and authorize the driver based on the veins in her finger.

Handling all the information exchange electronically also meant that the information available to each party in the transportation chain could be limited to only what was absolutely necessary. The goal is that as few people as possible should have access to as little information as possible about the goods and routes to reduce the risk of insider crimes. [9]

### 3.1.2 e-Freight

The Freight Logistics Action Plan is a policy initiative by the European Commission which purpose is to improve the efficiency and sustainability of freight transports in Europe. As a step to fulfilling these goals, the e-Freight project was launched as a joint initiative by 30 parties from 15 European countries [7]. Volvo is not a direct part of e-Freight, but has a close business relation with two involved companies, Stena Line and DSV. This collaboration has given good insight into the SITS project about what's going on within the work of e-Freight and especially the business case concerning the automated check-in at the Stena Line terminal. The project strives for paperless freight transport processes with a strong connection between electronic flow of information and the actual physical flow of goods. This will be done in practice by creating a framework to allow tracing of goods in real time by realizing the concepts of "single window" and creating the appropriate framework for the deployment of tracking and tracing technologies. "Single window" comprises the idea of a single entry point for information exchange between companies and authorities. It is a gateway for

standardized information that needs to be sent to meet mandatory reporting requirements. [8]

e-Freight's goal is to develop an overall framework in a distributed manner in the sense that each involved country will make their own implementation. The decision to make it distributed is based on the desire of the concerned parties to have control of their own data plus the complexity that would arise regarding the question of ownership of the data in a centralized system.

### 3.1.3 PEPPOL

PEPPOL [10] is an EU project with the aim of providing a standardized technological infrastructure for the procurement of contracts in the public sector across all European countries. One part of the project is the transport infrastructure that PEPPOL has developed called BusDox (Business Document Exchange Network) [11] which is based on a combination of W3C[1] and OASIS[2] standards. This transport infrastructure allows parties to exchange business messages in a secure and reliable manner. The e-Freight project will build a transport infrastructure with PEPPOL as a base, following the BusDox specifications. The SITS project will in turn use this for message exchange between parties, and then continue to build upon that by adding functionality to cover the entire flow of information.

[1] W3C is an international community where Member organizations, a full-time staff, and the public work together to develop Web standards.
[2] OASIS is a non-profit consortium that drives the development, convergence and adoption of open standards for the global information society.

## 3.2 Actors

A typical example of how a few actors connect to the system is shown in Figure 3.1. The authorities have each their own Access Point connected to a Single Window working as a single point of entry for the business to all involved authorities. The picture also shows that connections to the AP can be either automated by a strong coupling to the internal system (transport operator 1 and 3), or be manual by the use of a web interface (transport operator 2). The green communication links shows that transport operators have some form of short-range communication with the terminal operator when transports arrive at the gate.

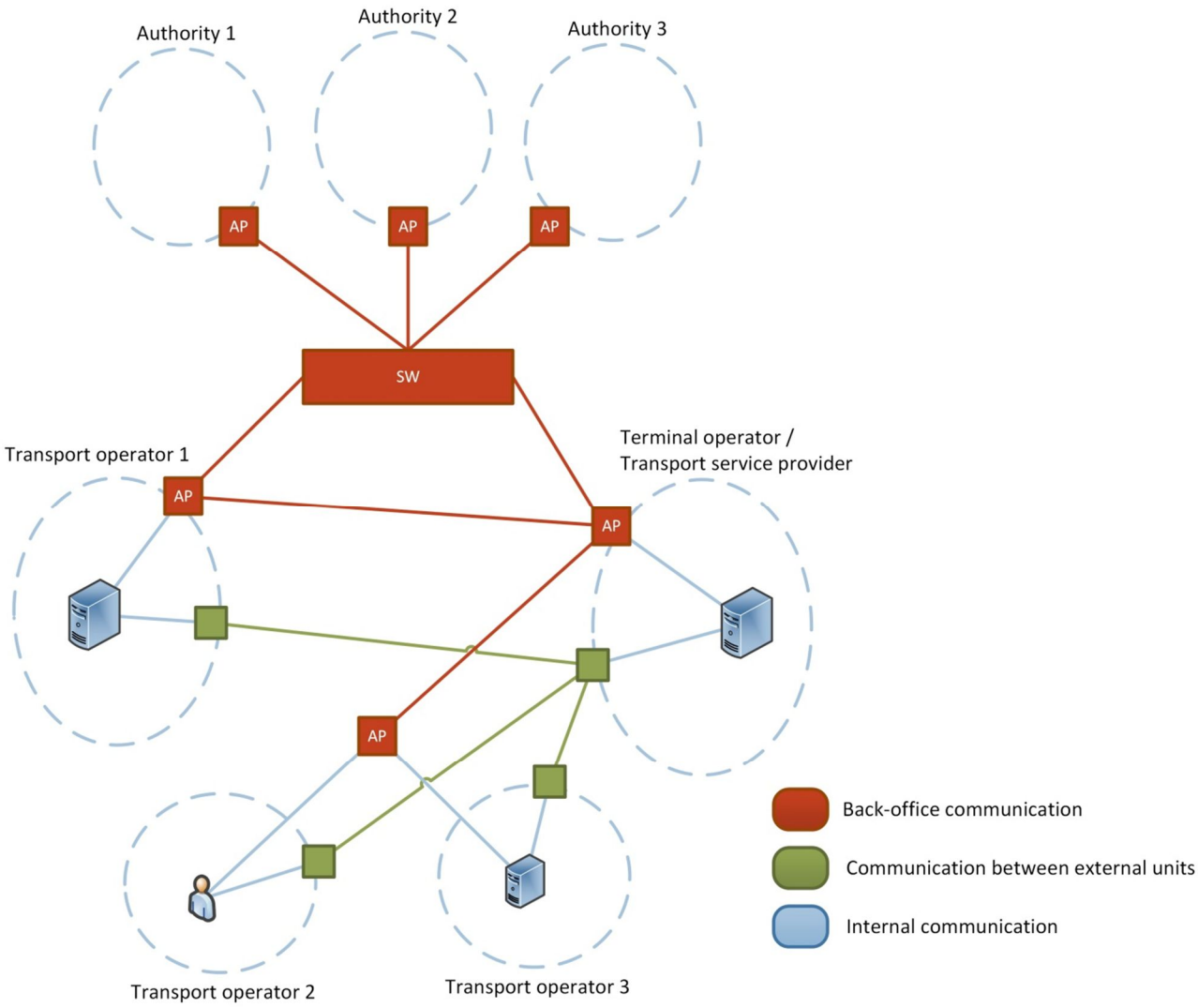


**Figure 3.1: Typical connections between actors and the system**

### 3.2.1 Transport operator

The transport operator is the one who physically carries the freight from one point to another. It is not mode specific so it can be by truck, boat, train or flight. A sufficiently large transport operator can set up his own access point through which back-office communication with other parties occurs whilst smaller companies have the capability to connect to a mutual access point.

### 3.2.2 Transport service provider

The transport service provider coordinates transports from one point to another. The physical transport is then taken care of by a transport operator and it might contain goods from several customers in a so called consolidate freight shipment, which the transport service provider sets up to maximize the effectiveness of the transport. This kind of operator is often big enough to host his own access point and have it connected to the internal system for planning and booking. The access point will then work as an external interface for its customers to the internal system.

### 3.2.3 Terminal operator

A terminal operator is one who manages a terminal. A terminal can be one of several types, for example a storehouse, a terminal for transshipment or a terminal for mode shift where goods loads from for example a truck onto a boat and vice versa. The terminal operators' internal system is strongly connected to the AP which receives information about transports, which are then being accessible down at the gate. When an incoming vehicle arrives at the terminal, the information is used to verify the vehicle and ensure that it is correct and has access to the terminal area. The actual verification is done with external units and there exist multiple possible implementations with different positive and negative aspects mainly concerning cost, efficiency and security.

### 3.2.4 Authority

There exist a wide variety of authorities with different areas of responsibility. The most important ones for the SITS project are customs, that checks incoming and outgoing transports for contraband and other illegal activities, authorities that are responsible for surveillance of for example dangerous goods, and emergency services that needs to act as fast and correct as possible in the event of accidents, especially when dangerous goods are involved.

Involved authorities are connected through their APs to the anomaly system and gets necessary information about deviating transports that needs attention.

## 3.3 Back-office communication

In the SITS framework all back-office communication between parties in the transport chain is done via a standardized interface: the access points from the e-Freight project. Whenever a company wishes to send some kind of message to another party the message is first delivered to the sender's AP, which will identify the AP of the receiver. The message is then delivered to the receiver's AP from which it is then delivered to the final recipient. The full details of how this works is described in the following subsection.

A company will be able to either set up and maintain their own private AP or, which will likely be the case for small and medium enterprises, they may negotiate for the services of an external AP maintained by a third party. Communication between the company and the AP can furthermore be coupled tightly with their internal business system to increase efficiency or it can be done manually through a web interface.

The framework will furthermore specify a message standard to be used in the communication between parties. The messages are in XML-format and based on a standard called Business Message Standard (BMS), which is developed by GS1 [16], an international non-profit organization.

### 3.3.1 Communication Infrastructure

In this section we provide a more detailed explanation of how messages are transmitted between different parties by describing some key parts of the network infrastructure in a BusDox system.

**SMP**
A Service Metadata Publisher (SMP) [27] is a server that contains the information necessary for a sender to be able to communicate with the receiver of a message. The sender first communicates with an SML (described in the next section) to find an SMP with information about the participant he wishes to communicate with. When the correct SMP has been identified the sender sends a request to it with the identifier of the other participant. The SMP then responds with the address of the AP that the receiver is registered with. Every AP needs to be registered with exactly one SMP.

**SML**
Service Metadata Locator (SML) [22] is the interface that allows for the discovery and management of SMPs.

The discovery interface uses Domain Name System (DNS) lookups where the SML servers act as DNS servers. When an AP wishes to find the SMP that contains the metadata about the AP that it is trying to contact, it performs an HTTP GET call with an address that is constructed using a standard format which contains, among other things, the business ID of the receiver. The GET response contains the metadata about the receiver. By using DNS, the system is able to function without a central server. The discovery flow is illustrated below in Figure 3.2.

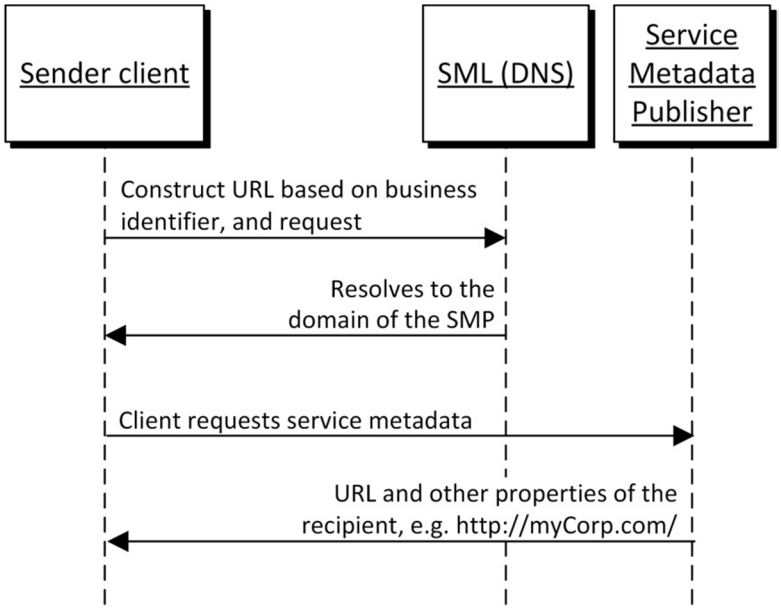


**Figure 3.2: The SML discovery process**

The management interface is used by the SMPs to create and update their associated DNS entries in a controlled way as depicted in Figure 3.3.

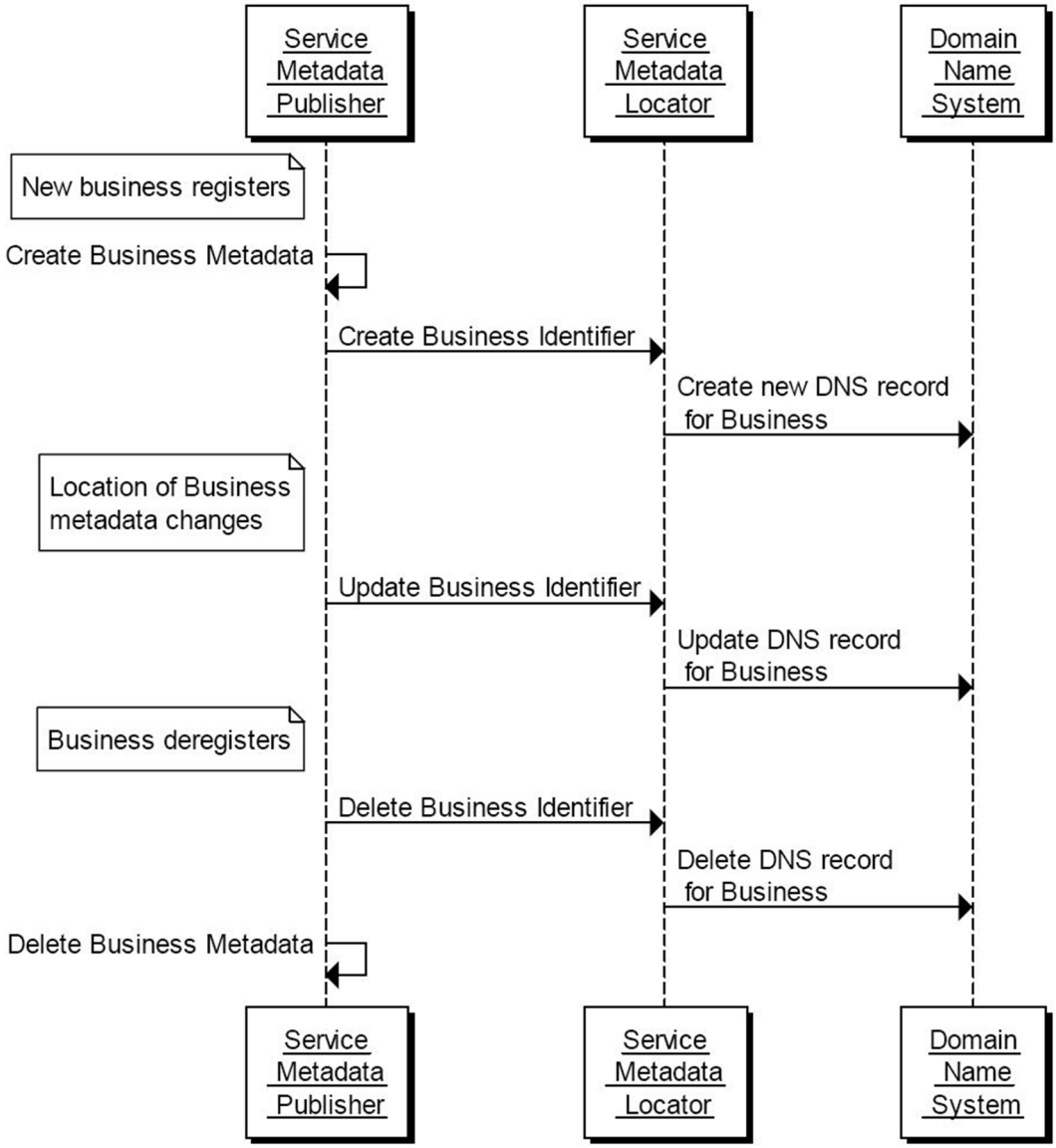


**Figure 3.3: Management of DNS entries**

### 3.3.2 End-to-end Communication Scenario

The sequence diagram in Figure 3.4 depicts a simplified view of a typical communication scenario. Here, Company 1 (C1) which is registered with AP1 wishes to communicate with Company 2 (C2) which is registered with AP2.

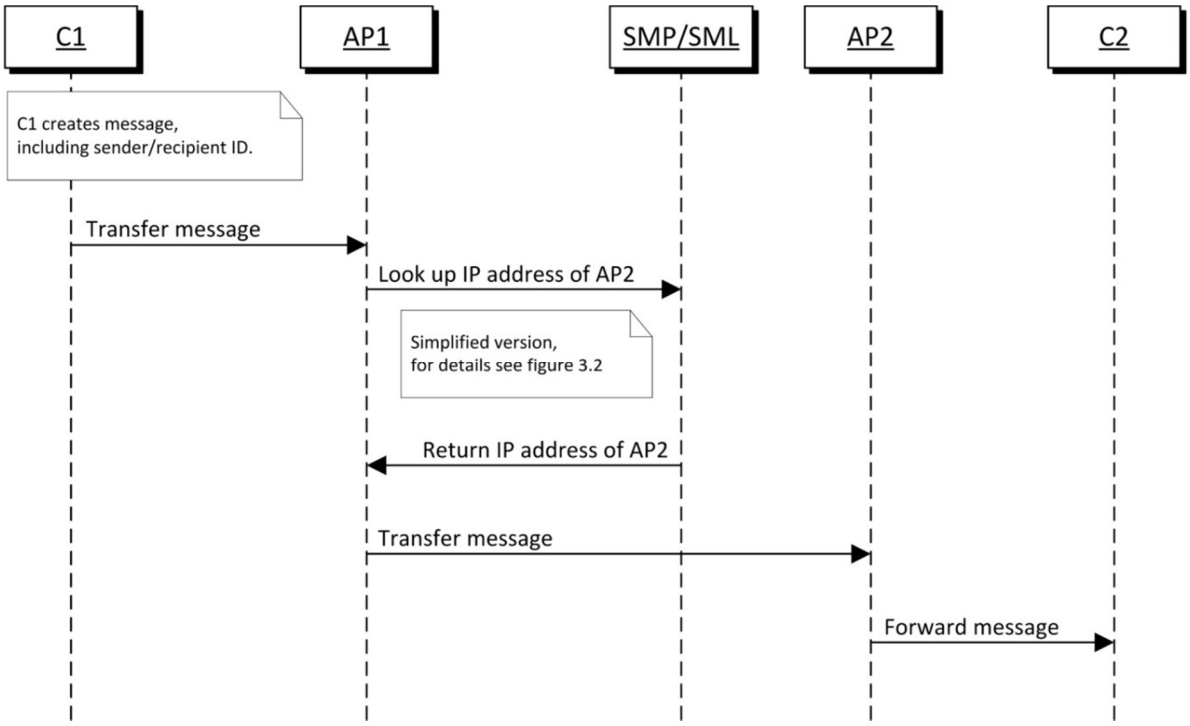


**Figure 3.4: An end-to-end communication scenario**

## 3.4 Short distance communication between external units

The SITS framework also encompasses short distance communication directly between two units, e.g. when a driver arrives at a gate to a terminal, data must be sent to the terminal operator for the transport to be authenticated and authorized before entering the terminal area. How this data is communicated and what specific hardware to use is left for the implementing actors to decide. But the lack of a mutual standard in this area increases the demands on the transport operators since they must adapt to multiple systems. This results in development cost and time and increased complexity for drivers needing to handle multiple systems.

Currently in the SITS project during development and for initial testing, the driver receives an SMS when he is on the way to the terminal, containing a six digit code that he has to use at the gate to be granted access. A few other possible techniques have been briefly reviewed. One of them involves a camera which would photograph the registration plate or another number or letter combination on the truck. This would infer no extra requirements for the haulers and all trucks could use the automatic check-in procedure assuming all required information needed for identification have been received beforehand by the terminal operator. There is a similar solution but instead of a number on the truck, a smartphone is utilized. A password in the form of a number or a kind of bar code is sent to the phone directly or indirectly from the terminal operator. The smartphone is then held in front of an optical reader at the gate for identification. Other solutions that have been looked into make use of RFID or NFC units to further simplify the check-in procedure by replacing the need of an SMS service and to have to enter a code at the gate. It also allows additional data to be transferred. The technology can be either placed in the vehicle or integrated into smaller handheld units. For example, there exists smartphones with NFC and it is also possible to integrate RFID into a small card that would be personal for each driver, similar to a VISA card. Unlike when a camera is used, these solutions impose a cost for the haulers, and if multiple

systems were to be used by different terminal operators, the increased complexity for the drivers would be evident. [9]

## 3.5 Automatic check-in

The automatic check-in concept is based on the Port pilot project 2008 and the e-Freight business case for booking ferries including pre check-in and check-in at a port terminal. To the e-Freight business case the SITS project adds direct communication to the vehicles and compared to Port pilot, SITS moves one step closer to a final solution by the use of existing message standards, harmonized communication processes and more mature technology. Another aspect of the system is that it is being developed in such a way that it shall be easy to transfer to other platforms and not be dependent on a specific hardware vendor.

The overall goal with this concept is to increase reliability of the information chain and increase the effectiveness in the transport chain. This is done by removing physical documents and use a fully IT based solution. The manual handling of documents will be decreased and by that make the system less prone to errors and at the same time make the information sent more complete. Automation is a key focus and will allow increased efficiency by a decreased manual handling of information and also by making the check-in more predictable (in a timely manner) since information can be verified before the vehicle has arrived physically. A more predictable process makes the transports more efficient by making it possible for better and more exact planning. Another positive aspect is that goods will be brought into a secured area faster which will decrease the likelihood of sabotage and theft.

The main idea is that before the goods have arrived, the terminal operator electronically receives all information necessary to automatically identify and verify that the incoming transport is authorized to access the terminal area. Which information is being communicated and how the verification actually is done, depends on the desired level of security. Standardization in this area is yet to be done.

Which parties that are being involved in the communication chain and in what way information travels differs depending on transport mission. In most cases, a transport is managed by a transport service provider who in turn hires a transport operator who physically transports the goods. However, it is possible that a single party acts as both transport operator and transport service provider. In this case the communication is done internally between the two roles. This does not affect the logics of the concept.

### 3.5.1 A Typical Check-in Scenario

A typical multimodal scenario where access to a terminal is needed is shown below from a communication aspect. In this example a buyer wants to transport goods where both a hauler and a shipping company are needed.

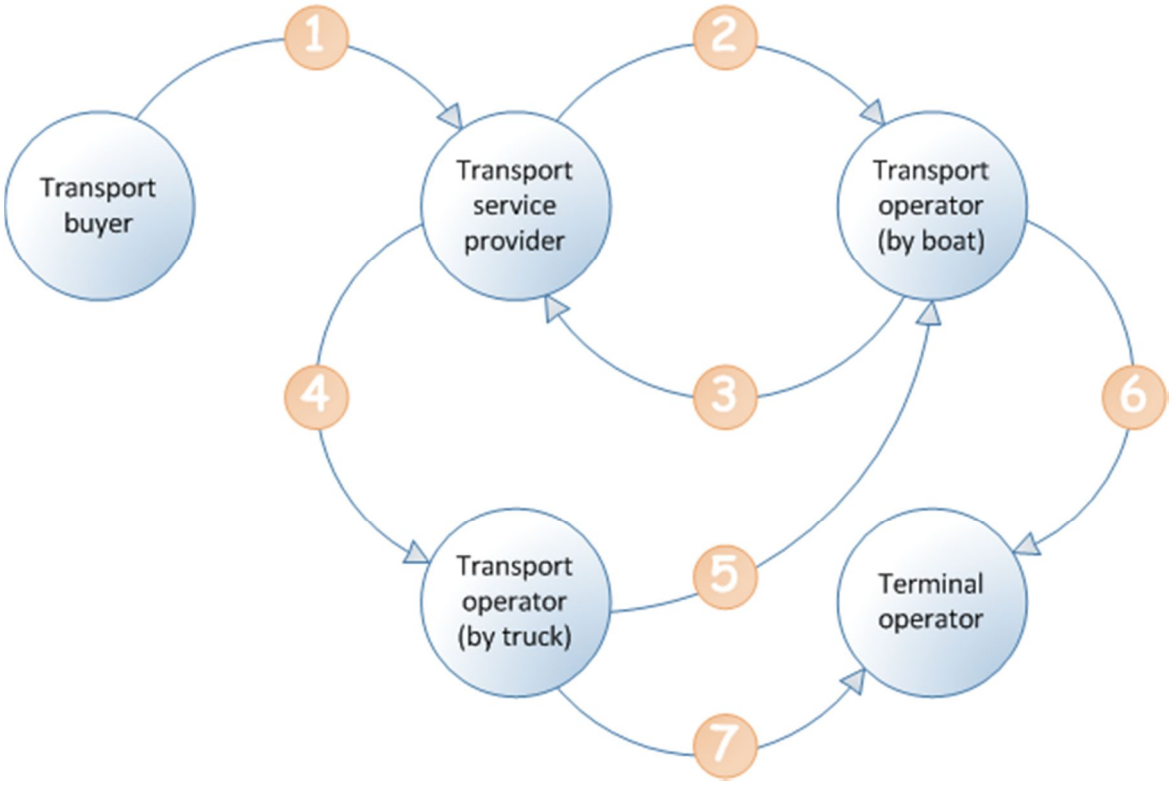


**Figure 3.5: Information flow in a typical check-in scenario**

1. The first step in the information chain is initiated by the buyer who places an order at the transport service provider who will take care of the transportation of goods for the buyer from one point to another. This step is not a part of this work since this information exchange is performed outside the scope of SITS.
2. The transport service provider then books a shipment and access to the terminal area. The booking includes information about the type of goods and time of arrival/departure. If the terminal cannot handle more incoming transports at the wanted time due to either high pressure at the gate from other incoming transports or because there is no room inside the terminal, the booking will not succeed.
3. The shipping company returns a booking number and possibly a valid time slot for when the terminal accepts the incoming transport. A time slot is optional and can be used for increased security since it makes it more difficult for a perpetrator to acquire access to the terminal area.
4. A booking request is sent from the transport service provider to the transport operator (the hauler). The information contains type of goods, time of arrival, pick up location, booking number and trailer id.
5. Information about the transport is sent to the transport operator (the shipping company) to be used for verification at a later step. The information can differ but generally includes booking number, trailer id, truck id and driver id. The shipping company replies with a release id used during the check in.
6. The shipping company forwards information about the transport together with the release number to the terminal operator.
7. The transport has arrived at the gate and uses short range communication to inform the terminal operator about who is coming. The terminal operator verifies the information against previously received data and if it matches, the transport is allowed into the terminal area. The shipping company is notified about the arrival and the goods will be shipped off [9].

# 4 Back-office Communication Security

In this chapter we describe the outcome of our analysis of the security of the back-office communication in the SITS project, i.e. the communication that is done via the access points. We start by stating the high-level security objectives of the system based upon discussions with key personnel at the companies involved in the SITS project. We then list some potential security threats against these objectives. We have categorized security objectives into the CIAAA (Confidentiality, Integrity, Availability, Authenticity, Accountability) model and the security threats follows the Threat, Vulnerability and Risk Analysis (TVRA) method [33] developed by the European Telecommunications Standards Institute (ETSI).

Finally we have evaluated the countermeasures against these threats that are provided by the BusDox specifications developed by PEPPOL and give some suggestions on how to further increase the security of the system.

## 4.1 Security Objectives

### 4.1.1 Confidentiality

Confidentiality is the concept of ensuring that data cannot be viewed by a party that is not authorized to do so. In the SITS framework, the messages that are exchanged between business partners may contain sensitive business information. It is thus crucial that the contents of such messages are not accessible to any other party than the ones involved in the transaction.

### 4.1.2 Integrity

Integrity means to ensure that data is protected from unauthorized modification, including data loss, data corruption or intentional modification by a malicious user. To ensure the integrity of the messages sent in the SITS framework there needs to be a mechanism to detect manipulation and corruption.

### 4.1.3 Availability

Availability means that data *should be accessible by an authorized party whenever it is needed. Since all information exchange in the SITS framework is done via the APs, it is crucial to the system that the downtime of APs as well as of the underlying system of discovering recipient APs is kept to a minimum, particularly during business hours.

### 4.1.4 Authentication

Authentication is the process of confirming the identity of a user or entity. In the SITS framework there are several places where an authentication scheme is needed. The receiving client needs to be able to verify the original sender of the message. Between a client and its AP there needs to be a mutual authentication scheme so that a client is not fooled into sending confidential messages to an entity impersonating the client's AP and so that the AP can safely vouch for the identity of the client in the message (see the section on SAML).

### 4.1.5 Non-repudiation

Accountability is a requirement that prevents the parties involved in a message exchange to later deny that the exchange did in fact take place. In other words, the

receiver cannot later deny that he has received the message and the sender cannot deny that he has sent it.

## 4.2 Security Threats

### 4.2.1 Interception

Eavesdropping is a breach of confidentiality that occurs when an unauthorized entity is able to monitor the communication exchange between parties and is able to glean information from this. Since the APs are not connected via some controlled private network but via the Internet, it should be an assumed that all messages that are sent can be intercepted by an attacker.

### 4.2.2 Manipulation

**Man-in-the-Middle Attack**
A modification attack is a general term for attacks where the attacker intercepts a message in transit and modifies the content of it. A man-in-the-middle (MITM) attack is an active form of eavesdropping. Instead of just passively listening in on the conversation, the attacker will intercept messages and modify or replace them with his own messages so that the communicating parties believe that they are talking directly to each other even though the communication is actually controlled by the attacker.

One way to achieve this could be with DNS cache poisoning. DNS servers are responsible for translating domain names into IP addresses. When a DNS server receives a query for a domain that it does not have a binding for, it queries another DNS server higher up in the hierarchy and caches the response for performance optimization. An attacker can intercept a query and respond with a false domain name - IP binding and thus poison the DNS cache so that all requests for that domain will be routed to an address of the attacker's choosing. Since the SML system uses DNS to resolve IP addresses of APs this attack could be used to direct traffic to a malicious server set up by an attacker.

**Spoofing**
A spoofing or (masquerading) attack is when an entity fakes its identity so that parties that are communicating with it will think that they are in fact communicating with someone else. An attacker could masquerade as an AP to intercept business sensitive information or as a business partner to send false information.

**Unauthorized access**
Unauthorized access is, as the name implies, when an unauthorized entity gains access to an entity. This access can then be used to perform other security breaches.

A malicious user that gained access to an AP can for example eavesdrop on or modify messages that are routed through it, forge messages that appear to be sent from its clients or harm the system in such a way that it ceases to function properly.

**Forgery**
Forgery is the act of fabricating information and claiming that the information was received from or sent to another entity. A malicious user could for example forge either a booking order or the confirmation of a booking from a business partner, thereby making it seem the two parties have entered into a business agreement even though the

other party has no knowledge of this. The other party could then be liable when he fails to uphold his part of this (forged) agreement.

### 4.2.3 Denial-of-Service

A denial-of-service (DoS) attack strives to make an asset, e.g. a server, unavailable to its users. This can be achieved in many ways. Some involve overloading the target network with traffic so that legitimate communication cannot get through, e.g. the “smurf” attack [25]. Others target vulnerabilities in the system, for example by sending malformed packets that the server does not know how to handle such as the “ping of death” attack [26]. A variation of DoS attacks are the distributed denial-of-service (DDoS) attacks. Instead of sending packets from a single attacking computer DDoS the attacker may use a large number of computers infected with malware (referred to as a Botnet) that flood a victim with traffic upon the request of the attacker.

## 4.3 Cryptography

In order to follow the discussions on the various security solutions presented in chapter 4.4 it is important that the reader has a basic understanding of some cryptographic principles. Therefore, we present here a short introduction to these subjects.

**Symmetric key cryptography**
In symmetric key cryptography the same key is used for encrypting and decrypting a message. This is the traditional way of doing cryptography and has been used for thousands of years (with substantial evolution of the actual encryption algorithms of course). The problem with symmetric key cryptography however, is that since both the sender and the receiver needs to have access to the same key, we need to be able to distribute that key in such a way that no third party will also be able gain access to it. The benefits of symmetric key cryptography are that it is relatively fast to perform encryption and decryption.

**Public key cryptography**
The problem of key distribution was solved in the 1970’s with the invention of public key cryptography. In public key cryptography, each party has two keys: one private key that is kept secret and a public key which may be distributed freely. When someone wishes to send a confidential message, it is encrypted with the receiver's public key that can then decrypt it using his private key, which only he knows. The downside of public key cryptography is that substantially longer keys are needed compared to symmetric key cryptography, which means that they are more computationally intensive and require more time. Because of this, public key cryptography is more commonly used to exchange a shared key between parties who can then use this key to encrypt their messages with a more efficient symmetric cryptographic algorithm.

**Digital signatures**
An interesting property of public key cryptography is that while encrypting something with the receiver's public key provides confidentiality, there is also a use for encrypting something with a private key. This does not provide confidentiality, since anyone can decrypt it using the corresponding public key, which should be widely available. However, since the receiver is able to decrypt it with the sender’s public key then he knows that it *must* have been encrypted with that person's private key and thus it provides a means of authenticating that the message was indeed sent by a specific person. A message encrypted with a private key is known as a digital signature.

**Certificates**
Even with public key cryptography one problem remains. A person's public key can be published to enable others to send confidential messages to him but the sending party cannot be certain that it was indeed the recipient who published the key and thus that he is the only one who will be able to decrypt it. Some malicious user may have published his own key in another person's name in order to decrypt messages intended for that person. The solution to this is public key certificates. A certificate consists of a public key plus the ID of the user it belongs to, which is then signed with the private key of a trusted third party, called a certificate authority (CA). If one is able to decrypt a certificate with the CA's public key (which needs to have been obtained in some secure manner) then I know the certificate was indeed issued by the CA. The CA also maintains a certificate revocation list (CRL) which contains, as the name implies, a list of all of the certificates it has issued which are no longer valid, e.g. because they have expired or been compromised in some way. A client can thus check a certificate against this list to be sure that it has not been revoked.

**Hashes**
Another form of cryptographic functions are hash functions. A hash function takes a message and converts it into a fixed-size string known as a "message digest" with certain properties:

- The message digest can be computed very quickly.
- It is computationally infeasible to find a message with a given hash.
- It is computationally infeasible to find two messages with the same hash.

Hash functions are commonly used in combination with a secret key to ensure message integrity in the form of a hash-based message authentication code (HMAC). The sender of a message calculates an HMAC as a combination of the message and a secret key and then appends HMAC to the message. The receiver has previously obtained the same secret key and can therefore do the same calculations upon message acquiring. The received HMAC and the calculated HMAC is then compared and if they both match, the message integrity has not been compromised.

Another common use of hash functions are in storing passwords. Rather than storing passwords in cleartext they are encrypted using a hash function and the hash value is stored instead. When a user later enters a password this too is hashed and the result is compared with the stored value. To further increase security a random value, known as a salt, can be added to the password before it is hashed. The salt is then stored in cleartext together with the hash digest. The purpose of this is to make it even harder for an attacker who has gained access to a password database to recover any passwords. If a hash function is used without a salt then the attacker can generate a dictionary of hashed words and compare this to the hashed passwords in the database. By adding a salt to the function the attacker would have to generate a new dictionary for every salt.

Hashes can also be used to create more efficient digital signatures by having the signee sign the message digest instead of the entire message.

## 4.4 Security Solutions

The communication methods used in the BUSDOX infrastructure are specified in two different profiles: the Lightweight Message Exchange (LIME) profile [4] for client-to-AP communication and the Secure Trusted Asynchronous Reliable Transport (START)

profile [5] for AP-to-AP communication. This section will describe the used security features in these two profiles.

### 4.4.1 Transport Layer Security

Both the LIME profile and the START profile include the Transport Layer Security (TLS) protocol as the means to ensure the confidentiality and integrity of the messages exchanged. TLS and its predecessor SSL (Secure Socket Layer) are widely used on the Internet to secure communications, e.g. when providing login credentials or credit card information on web pages.

At the start of a session, the TLS handshake protocol negotiates which cipher suite to use, i.e. the combination of a public key algorithm used for key exchange, an encryption algorithm used for the message stream and a MAC (Message Authentication Code) algorithm. The decision on which cipher suite to use is determined by the client and server exchanging a list of supported cipher suites, listed in order of preference, and the most preferred suite that both parties support is then agreed upon. Consequently, if a client feels that a particular cipher suite does not provide the level of security that the communication exchange warrants then that cipher suite should be removed from the list of supported suites.

It should be noted that as a consequence of the communication infrastructure that is used in this project, TLS cannot be used to provide end-to-end security. Since messages are relayed through access points, one TLS connection will be created between the sending client and its AP, another between the senders AP and the receivers AP, and a third one between the receivers AP and the receiving client as illustrated in the figure below. What this means is that should a malicious user gain access to a message at an AP then he can both read and modify it unless some form of protection is implemented at a higher level. This highlights the importance of being able to trust the provider of the AP service at both ends of the communication in the service is outsourced to a third party.

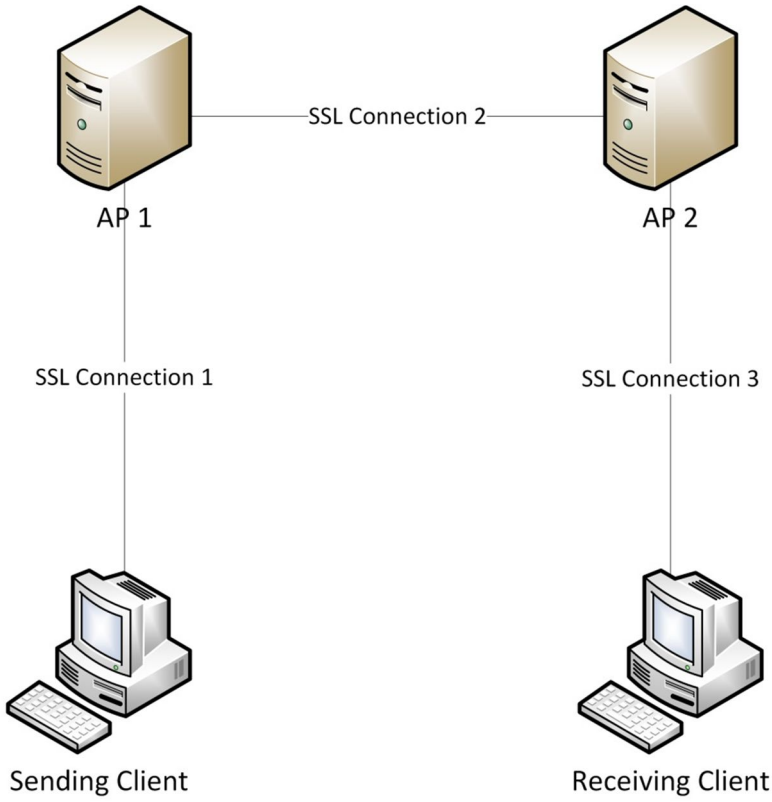


**Figure 4.1: SSL connections in an end-to-end communication scenario**

### 4.4.2 HTTP Basic Authentication

The LIME profile uses HTTP Basic Authentication [1] in order to authenticate the client to the AP. The scheme is the simplest form of password authentication, i.e. the AP asks for a username and password and the client responds with its credentials.

The biggest weakness in the Basic Authentication scheme is the fact that the username and password are transmitted in cleartext so that an attacker could steal the credentials by simply eavesdropping on the connection. However, since all HTTP traffic is protected by TLS/SSL in accordance with the LIME profile, this weakness should be of no concern.

### 4.4.3 PKI

In order to manage the use of certificates in a network a public-key infrastructure (PKI) is needed. IETF's Internet Security Glossary defines a PKI as "the set of hardware, software, people, policies, and procedures needed to create, manage, store, distribute, and revoke digital certificates based on asymmetric cryptography" [2].

An important part of a PKI is the CA. Since every AP in the system needs a certificate that is signed by the CA this offers a way to enforce security policies. In the PEPPOL system every AP provider needs to sign the PEPPOL AP Provider Agreement [3] which regulates the providers' responsibilities, requirements and liability. A similar agreement will likely be drawn up for the e-Freight project.

### 4.4.4 SAML

In order to authenticate the original sender of a message to the end recipient the START profile specifies the use of the Security Assertion Markup Language (SAML) 2.0 [6].

SAML is a standard that allows for the exchange of authentication information between different security domains. What this means is that when the sender's AP forwards a message on behalf of its client it includes information about the identity of the sender and details of the sender's authentication information. This can be done in one of two different ways. The first way is that the AP itself authenticates the client; this is known as "sender-vouches" assertion. The AP encloses the identity of the sender and the method used to authenticate him and signs this information, thereby "vouching" for its validity. The receiver thus needs to trust the sender AP. The other way is called "holder-of-key" assertion. In this method the sender's AP requests for a trusted third to sign the SAML assertion instead of signing it itself.

### 4.4.5 DNSSEC

To prevent an attacker from performing DNS poisoning the SML specification [22] recommends (but does not require) using DNSSEC which provides authentication of DNS lookups by having the result digitally signed by a trusted third party.

However, even if DNSSEC is not used, protection from spoofing attacks can still be ensured as long as the AP is required to authenticate itself to the client. If a client has been redirected to a malicious AP then the AP will fail to authenticate itself before any confidential messages are sent.

### 4.4.6 Protecting Against DoS Attacks

DoS attacks that rely on exploiting bugs in the target software can usually be combated by simply patching the software to remove the vulnerability when they are detected.

Attacks that flood the target system with traffic, particularly DDoS attacks, are however notoriously hard to protect against. DDoS attacks of limited size can be prepared for by provisioning for a higher amount of traffic than the legitimate business generates [32] and some forms can be mitigated by having a properly configured router [31]. However, when under a large scale DDoS attacks the only option may be to work with the ISP to try to stop the traffic upstream.

# 5 RFID Security Issues

This chapter describes security issues regarding RFID technology since its use in SITS is under consideration. Highlighting concerned areas with this technology can help make a more accurate decision in how the short range communication in SITS in practice will be realized and eventually move towards standardization for this kind of application within the business. There is also a suggestion to use NFC-enabled smartphones and since NFC is a subset of the RFID specification, the issues in this chapter are also relevant for the NFC technology [12].

## 5.1 Description of RFID

An RFID system comprises three distinct parts; tags, readers and a back-end system. A tag and a reader communicate wirelessly with the use of radio-frequency electromagnetic fields. The data sent to the reader is forwarded to the back-end system during for example identification or authentication of the tag. A tag can in turn be divided into three different categories; passive, semi-passive and active. A passive tag is completely powered by the reader from the radio energy that is being emitted. The available resources in forms of for example cryptographic functions on a passive tag are highly restricted due to the limited power consumption. A semi-passive tag has an internal battery used for computation but uses power from a reader for transmission, while an active tag has an internal power source for both computations and transmission. These two kinds of tags can be equipped with significantly more sophisticated functions compared to passive tags. On the downside, this results in increased size and higher costs. An active tag also has a larger range as the signal power from the tag is not dependent on the reader and can therefore be stronger, compared to passive and semi-passive tags which receive a very limited amount of power and cannot afford the same level of signal strength [13].

RFID is specified for a wide variety of frequencies and the operation range can therefore differ quite a lot between different RFID systems. NFC however is just a subset of RFID and operates at exactly 13.56 MHz and with a range of typically 10 cm [14]. The operational range is not totally definite because it is affected by factors such as transmit power, antenna, surroundings etc. which might be adjusted by a malicious user to his favor [15]. The choice of having a short operation range is a conscious decision since the technology is aimed for use in smartphones for applications such as digital wallets and for bootstrapping other more competent wireless communications [17]. A short range is wanted for these types of applications to make sure the communication is only carried out between the two physical devices that it is intended for. This makes the technology more secure when it comes to relay attacks and eavesdropping as the attacker must position himself closer to the devices to be able to listen to and possibly alter the communication stream.

From a security point of view an RFID system can be divided into the three different parts: *Edge Hardware Layer*, *Communication Layer* and *Back-end Layer*. The RFID *Edge Hardware* consists of tags and readers and is concerned with the security of these physical units. The *Communication Layer* deals with the exchange of information between readers and tags. The last part, the *Back-end Layer*, is the back-end system that is connected to RFID readers and often contains a database, a web server and some kind of middleware.

The rest of this chapter describes prominent attacks against these three areas. The analysis will is based on the classic CIA model that categorizes attacks depending on if they compromise Confidentiality, Integrity or Availability. The chapter ends with a presentation of available and upcoming countermeasures to identified attacks.

## 5.2 Edge Hardware Layer

The Edge Hardware Layer involves readers and tags and is concerned with direct attacks against these units at a physical level.

### 5.2.1 Confidentiality

The confidentiality at the Edge Hardware Layer can be compromised by a so called side-channel attack. Information about power consumption, signals, radiation and fluctuation in timing delays may be obtained. This information can be used to reveal secret data such as cryptographic keys or seeds used in cryptographic operations. There exist several documented attacks in this area and some of them use Simple Power Analysis (SPA) and Differential Power Analysis (DPA) which are non-invasive and can be used against tamper-resistant devices. Both of these techniques measure variations of the power consumption. SPA is simpler and involves analyzing time-resolved electric current measurements directly, while DPA is not as sensitive to noise and involves a more advanced statistical analysis of the power consumption [18]. Another related technique is Electromagnetic Analysis (EMA) that instead of measuring the current, electromagnetic field variations are measured to uncover cryptographic secrets [21]. Data resulting from an EMA is not necessarily the same as from a Power Analysis and it can in fact be used to reveal power analysis countermeasures and neutralize them [20]. It is also possible to measure the time between certain operations and secrets may then be discovered based on the noted delays.

In addition to side-channel attacks, confidentiality is also threatened by *fault attacks*. These are conducted by producing errors in the operation of a device by exposing it to abnormal environmental conditions like heat, cold, electromagnetic radiation or high voltage levels. Sensitive information can then be retrieved as a result of this either through normal communication channels, for example the one used between a reader and a tag, or it can be retrieved through side-channels, for example by monitoring the power consumption.

In contrast to the so far mentioned attacks, an attacker can employ a more hands-on method and physically take apart the device to retrieve interesting data therein. Physical tampering can also be done by changing or damage specific parts in the hardware. For example by changing targeted cells in ROM to change the behavior or seed of a cipher [18].

### 5.2.2 Integrity

Integrity is a security aspect which can be divided into data integrity and system integrity [23]. Data integrity at the edge hardware layer involves physical data modification and would typically result in an invalid tag. This kind of attack does not lead to unauthorized access to the terminal area which is the most important asset that needs protection, but it does affect the availability of the system since the driver can no longer use the automated check in.

System integrity on the other hand is concerned with correct operation and handling of information in the system. Tag cloning and spoofing are the most prominent attacks in

this area. With tag cloning the goal is to make an exact copy of a legitimate tag, concerning both the data on the card and possibly even the physical appearance. If no security mechanisms are present this kind of attack can easily be deployed by simply reading all data on the card and transfer it to another tag. The required knowledge to launch that kind of attacks is low and the needed equipment such as blank tags is freely available. But if any form of authentication or encryption is used there will be data on the card that is never sent and revealed such as a secret password. For an adversary to clone a card of that type he will need to break the encryption which is highly unlikely when dealing with standardized protocols. Closely related to tag cloning is spoofing which instead of copying a tag, only simulates one. This can be done by first eavesdropping a transmission from a tag and then at a later time retransmit the same data using a reader connected to a computer, a so-called replay attack. Contrary to attacks against data integrity, both tag cloning and spoofing opens up for unauthorized access to the terminal and should definitely not be underestimated [18].

### 5.2.3 Availability

The availability of the system for authorized users is threatened by physical destruction and other kinds of sabotage. The most obvious threats are vandalism of the reader and the risk of it being stolen. Also a tag can be rendered unusable for example by the use of the RFID Zapper. It is a device that loads a capacitor with high voltage and then generates a strong electromagnetic field with a coil that can blow out parts in a passive RFID tag if it is close enough [19].

The Edge Hardware can also be temporarily disabled as a result of extreme environmental conditions such as ice or water covering the reader or tag. Another more technical oriented attack that might be possible is exhaustion of protocol resources. The possibility to deplete protocol resources arises from the fact that certain protocols have limits on how many times a tag can be read or how many unsuccessful reads are allowed before the tag is turned inactive. There also exist protocols that use time or counters with a fixed maximum value. An attacker can take advantage of this and deliberately make the counter rapidly increase and when it has reached its maximum value the tag can no longer be used. Attacks focusing on consuming the battery power of an active tag might also be possible by setting up bogus communication links with the tag and by that shortening the lifespan of the tag [18].

## 5.3 Communication Layer

The Communication Layer focuses on the information exchange between readers and tags. Eavesdropping, man-in-the-middle attacks and crypto attacks are typical threats against the RFID communication.

### 5.3.1 Confidentiality

The most obvious threat against confidentiality is *Eavesdropping*, as is the case with all wireless communication and RFID is no exception. However, RFID typically operates at a much smaller distance compared to Wi-Fi because of the small form factor which demands low power consumption and small antennas. But even if the range is very limited for a legitimate reader and tag, it does not mean that the same limit is imposed on an attacker. Special long range directional antennas exists that can be used to capture RFID communication at a far greater distance than the operational range of the system. Even if the communication channel is encrypted, secret information may be revealed through traffic analysis and used for more sophisticated attacks.

Another related concern is *Unauthorized Tag Reading*. If no authentication mechanism is implemented in the tag, the only obstacle for an attacker to be able to read the tag is to get in range.

If the security of the communication layer is improved with encryption and authentication mechanisms, *Cryptographic attacks* become the focus. The main things an attacker might seek to reveal are a possibly shared key, temporary encryption keys and secret readable information stored on the tag. Examples of attacks that might be possible are the use of brute force to expose passwords or plaintext, chosen ciphertext or known plaintext attacks against the cipher and attacks against hashes such as preimage and collision attacks [18].

### 5.3.2 Integrity

For the automated check-in system, the most important integrity attacks against the communication channel are replay and relay attacks. If one of those were successfully launched by an attacker, unauthorized access into the terminal would be gained.

Replay attacks are basically done in two steps. First, data communication is eavesdropped when a legitimate driver checks in. Then at a later time the same data is resent to the reader at the terminal. If the system accepts the data, the adversary can enter the terminal. However, this can easily be thwarted with the use of some kind of challenge/response mechanism involving random numbers or by using a time stamp. It would also suffice to let the driver access the terminal only once with the given data that is presented to the reader. In this case the attacker will be denied access when sending data to the reader that has been received earlier.

A far more serious and complex attack is the man-in-the-middle attack. Here an attacker places himself between a tag and a reader and communicates with both of them in a way that they believe they are talking to each other even if they are far apart. Even if authentication and encryption mechanisms are implemented, this kind of attack can still be done. The attacker does not need to know any secret keys or authentication credentials because any kind of challenge from the reader is relayed to the tag which computes a correct response which is then sent back to the reader. A possible scenario involves a driver that has not yet checked in, an adversary at the gate with a special tag connected to a computer, a reader under the attacker's control that is close to the legitimate driver. The attacker's reader and special tag are connected through the Internet. Communication is initiated by the attacker at the gate and he receives a challenge for authentication. The challenge is sent to the bogus reader and presented to the legitimate tag which computes a response. The response is sent back to the fake tag and forwarded to the reader at the gate. The reader now believes that the legitimate driver is at the gate and wants to check in and therefore grants access to the terminal area [18].

### 5.3.3 Availability

The communication channel between a reader and a tag can easily be disrupted by an attacker and as a result cause a denial of service. One way is to use electromagnetic interference that can make the communication unstable and unable to operate correctly. Another way is to use a so called Blocker Tag. It was developed as a result of the introduction of embedded RFID tags into products such as cloths. These embedded tags are replacing the traditional barcodes and can be used during check-out before paying and also to improve inventory management. The problem is that an adversary can monitor people's shopping habits which are an infringement of personal integrity. What the Blocker Tag does is to simulate a wide variety of products and when queried by a

reader, responds with unique ids of several tags. This behavior makes it difficult for an adversary to distinguish which ids that are real and which that are not. But the problem with the Blocker Tag is that it can also be used in a mischievous way. For example, imagine that it is placed closely to a reader. This would render that reader completely useless as it would be flooded with ids where only a small percentage is real. But this threat only arises for simple tags which lack any form of authentication mechanism.

## 5.4 Back-end Layer

The Back-end system ties parts of the entire system together. It is generally composed of a server, software and some kind of middleware. The software contains a web service that allows information to be received from the main booking system. When a booking is done for access to the terminal, the necessary information needed for authorization during the check-in is sent to the back-end server and stored in a database. Later when the transport operator arrives at the gate, the RFID reader communicates with the server through the middleware that creates and sends a query to the server based on the data received. Data from the RFID tag is looked up in the database before authorizing the driver into the terminal area.

Attacks against the Back-end Layer have two distinct entry points. One of them is through networked connections. For example sharing or linking of databases with other parties, internet connection or usage of an Object Name Service (ONS). The ONS mechanism is similar to the Domain Name System (DNS) and is used to convert an Electronic Product Code (EPC) to an address pointing to a locally or globally stored file containing information about the product. The process starts with the middleware receiving the EPC sent from the tag. DNS is then in fact used for the actual lookup and the middleware receives the file after proper authentication. The other entry point for attacks is the data coming in from the edge hardware. Only these kinds of attacks will be considered in this chapter since network based attacks are not unique for RFID and the standard IT security policy used within the company should handle this [18].

### 5.4.1 Confidentiality

The most important information stored at the back-end that needs protection is about which driver will be arriving at what time and also the cryptographic keys used in different situations. A compromise of either of these, and especially the cryptographic keys, can have devastating results. An attacker might be able to access or impersonate tags and readers at will if an attack can be launched via the edge hardware.

### 5.4.2 Integrity

The integrity at the Back-end Layer is mainly threatened by code injection for example through buffer overflow attacks when it comes to attacks derived from the edge hardware. For code injection to be possible, a vulnerability has to be used in the middleware. And since middleware applications often are written in multiple scripting languages, it is not unlikely that security holes exist. Buffer overflow attacks are based on sending more data than what is being expected and if no care has been taken, data might overwrite other parts of the memory than what is intended. RFID tags have a very limited memory and can therefore not send as much data as an adversary might want. There are other ways to do it though by the use of devices with more memory such as smart cards or devices that can emulate multiple RFID tags. If the in-data from the tag is not checked properly, hacker attacks such as trojans, worms, viruses, SQL-injection etc. can potentially be propagated through the reader to the back-end [35].

### 5.4.3 Availability

A Denial of Service attack against the back-end will render the check-in system totally unusable if no backup systems are available. A DoS attack from the edge-hardware can be in the form of flooding or spamming messages from a tag to the reader. If no detection is done at the readers these messages will be forwarded to the back-end server and potentially make it unavailable.

## 5.5 Countermeasures

The use of encryption for the communication stream is a good way to counter eavesdropping as it makes all data sent between a reader and a tag useless for an adversary. A technique that goes hand in hand with encryption is authentication which efficiently thwarts unauthorized tag reading. These are mainly solutions for active and semi-passive tags since the power in a passive tag usually isn't enough for advanced computations. However, upcoming protocols do exist for passive tags that offer cryptographic functions at low power costs [28] [29]. Still, it is worth to keep in mind that new protocols might contain security flaws due to the lack of extensive security testing. The general idea in this area is to use strong, published and well-known cryptographic algorithms. History has shown that security by obscurity shall not be practiced as the internal design of the system eventually will leak out. The closest related example is the RFID system used in the Dutch transportation system. Here the proprietary CRYPTO1 encryption algorithm was broken and as a result, the cryptographic keys could relatively easily be retrieved [30].

Another concern with encryption and authentication in passive tags is the achieved level of security. It is often desirable to have long keys and a challenge-response protocol for authentication with a dependence between the challenge and the response with the use of for example timestamps, a counter or a nonce. This might be difficult to accomplish with limited computation resources.

Man-in-the-middle attacks are still left unaffected by authentication and encryption in the sense that it is still possible to relay messages between legitimate tags and readers. The problem becomes evident when considering the scenario in which the adversary is at the gate and the legitimate driver has not yet checked-in. But such an attack only succeeds if an attacker is able to read the legitimate tag. The short operational range of RFID certainly enhances security but does not make it impossible for an adversary. A more definite countermeasure is to only let the tag be read when the driver is at the gate. This can easily be achieved for active tags. For example when using NFC in a smartphone, the user interface can ask for input from the user to approve the communication. However, passive tags do not come with this kind of on/off switch. They do not have as long range as active tags though, but are still vulnerable. One countermeasure for passive tags is to use a shielding cover when not being used to reflect radio waves and make the tag unreadable. A different point of view for countermeasures against relay attacks is to use detection instead of direct protection. It is possible for a reader to detect when signals from a tag is being relayed by the fact that a relayed signal has a longer round-trip time compared to when a tag is in direct contact with the reader. The solution for this is called Distance-bounding protocol and there exist several protocols of this kind in the literature. They are all based on the same general idea. A reader and a tag share a secret key that together with a hash function and nonces from both the tag and reader generate two different binary sequences, S0 and S1, at both the reader and the tag. This is followed by a number of challenges sent by the reader in form of zeroes and ones. When received by the tag it replies with the

corresponding bit from either S0 if a zero was sent or from S1 if it was one. The reader then compares the responses with S0 and S1. During the part where challenges and responses are exchanged, the reader keeps track of how long time it takes. If the responses are correct and the time taken is under a specified threshold, the tag will be validated by the protocol and considered close to the reader. The main problem is the computational delay at the tag during the challenge-response part. For the reader to be able to calculate the distance as accurately as possible the delay at the tag must be extremely low to have as little effect as possible. This is the reason S0 and S1 is computed beforehand and then just accessed when needed. To further decrease the computation time at the tag, S0 and S1 are stored in two shift registers. When the tag gets a challenge, the correct register is selected which directly returns the first bit asynchronously and does not wait for any clock cycle. The first bit in both registers is then discarded at the same time. This kind of protocol does not necessarily put any demand on the computational power except the need for the shift registers. This means that it can be implemented in both active and passive tags [34].

A more general security solution involves the use of a timing window which only allows the driver to enter the terminal within a given timeframe. This will make it harder for an attacker to gain unauthorized access and with the combination of only letting the driver enter the terminal once during his current transport mission, faster detection of the adversary will be possible as the driver will react when he cannot enter. It might also be reasonable to make use of the transport information system in those trucks it is available in to confirm the geographical position of the truck before granting it access into the terminal, thwarting relay and replay attacks.

The back-end system is mainly threatened by common attacks for networked systems. These attacks are often already countered in large companies by the use of firewalls and intrusion detection systems. It would therefore suffice to place the back-end system behind these and possibly add some rules in the firewalls and IDS. Another type of attack that needs to be countered is code injection originating from an RFID tag. That can be done using the same kind of techniques used in web servers when handling data from web forms supplied by users. Checking for illegal characters, data length, or something else that’s not expected.

The availability of the system is an important feature although the lack of it does not allow unauthorized access into the terminal. The most serious attack against the availability is active interference in form of radio waves. A possible solution is to use shielding material around the reader to only allow signals from a very limited angle.

# 6 Results

We found that the PEPPOL specification use well-established security standards in order to ensure secure communications. However, the central roles that the APs have in a BusDox infrastructure make it extremely crucial that they be protected from attacks from malicious users. Therefore, we have proposed some security requirements that should be followed by an AP provider. These should preferably be specified in the agreement that the provider signs before he is issued a certificate. The requirements are:

- Anti-virus software should be installed on the server that hosts an AP and the virus definitions should be updated regularly to protect against new viruses.
- The passwords that are stored should be in a hashed and salted form to limit the damage should an attacker gain access to the password database.
- Enforce the use of strong passwords to prevent the use of Rainbow tables for quick retrieval.
- A well-configured firewall should be set up to protect the AP.
- Port scanning tools should be used to identify and close unused open ports in the AP.
- The OS and software on the server should at all times be kept up-to-date with security patches.
- Keep the number of services installed on the AP to a minimum by only installing needed services.
- An intrusion detection system should be used and monitored.

We can conclude that the use of RFID at the gate is a good way to handle authentication as the process for the drivers is simplified at the same time security is strengthened. We consider attacks against availability to be a low risk even if attacks involving for example radio frequency interference can easily be deployed. The only motive for an attacker would be to delay passage with no direct personal gain. The risk of getting caught is always present since an adversary could easily be located by measuring signal strength or direction of the interfering signal. Basic attacks such as eavesdropping and tag cloning can easily be thwarted by implementing standard security in form of encryption and authentication, but this still leaves the system vulnerable to man-in-the-middle-attacks. There exists a few variants but the most potent type of MITM-attack allows an adversary to enter the terminal area by relaying the communication between a legitimate tag and the reader. The communication to the legitimate reader and between the attackers' own devices do not impose any advanced hindrances although it requires both expertise and possibly expensive equipment. The real challenge for an attacker to succeed with such an attack is to be able to read the legitimate tag. It is therefore important that the tag can only be read when the driver chooses to. This makes a tag installed into the vehicle not suitable if it can be read by anyone, anytime, even if encryption is applied. NFC in smartphones on the other hand is a good choice. Even a personal driver's card is a good solution if some kind of physical protection is used, for example a casing made of insulating materials to thwart unwanted reads.

Other solutions that are viable to increase security in combination with RFID are the use of timeslots and geofencing. To predict when the driver is allowed to arrive at the gate makes timing a crucial factor for an adversary. Geofencing can be utilized through the in-vehicle computer and be used to strengthen the authentication process by having an extra channel of information ensuring that the truck is really at the gate. There are also protocols supporting distance-bounding, meaning that replies must arrive quickly and there is not time to forward it to an external system and act as a man in-the-middle.

# 7 Conclusions and future work

In this report we have analyzed the security of two parts of the SITS project: the back-office communication and the RFID-based system for automatic check-in at terminals.

For the back-office communication we established the security objectives that should hold for the system and discussed some possible threats to these objectives. Since the communication infrastructure that SITS will use will be based on the one used in the PEPPOL project, we have analyzed the security countermeasures that this project uses to combat these threats.

When researching RFID security, most literature typically focuses on environments such as stores and hospitals which are completely different compared to the environment considered in the SITS project. Tags attached to every item in a store needs to be low cost and not easily removable and in a hospital where tags represent equipment or medicine, the data in the tags must be highly accurate. In the SITS project however, the main threats against tags are theft and tag cloning while manipulation of tags is not as important. Tag-manipulation might affect the availability of the system though, but the primary asset to protect is the terminal area.

Our opinion is that SITS is a reasonably secure system with respect to external threats. Developers in both PEPPOL and e-Freight have had security in mind during the development process. Also, RFID can safely be used but with caution against proprietary algorithms and new protocols because of the unknown security flaws it might contain. What kind of tag to use is a matter of trade-offs. Passive tags are typically harder for an adversary to read because of the limited transmission range while an active tag can comprise more advanced security. Man-in-the middle attacks must also be considered when RFID cards are used.

As the final design of the SITS project has not yet been deployed there has been no way for us to do practical tests of system security. We therefore recommend that once the system is ready, that a more hands-on analysis be performed. The APs are connected to the internet and shall therefore be tested against common network penetration techniques. Interfaces, both internal between APs and external with human users or other systems, need to be checked for vulnerabilities to find possible faults in the protocols. In addition, the subsystem handling the short-range communication would also benefit from a thorough hands-on security inspection.

# 8 References


[1] J. Franks, P. Hallam-Baker, J. Hostetler, S. Lawrence, P. Leach, A. Luotonen and S. L, "HTTP Authentication: Basic and Digest Access Authentication," 1999. [Online]. Available: http://www.ietf.org/rfc/rfc2617.txt. [Accessed 4 July 2012].

[2] R. Shirey, "Internet Security Glossary, Version 2," 2007. [Online]. Available: http://datatracker.ietf.org/doc/rfc4949/. [Accessed 4 July 2012].

[3] PEPPOL, "PEPPOL AP Provider Agreement," 16 May 2011. [Online]. Available: http://www.peppol.eu/about_peppol/PEPPOLAPProviderAgreement_v2p0r1_20110520.pdf. [Accessed 4 July 2012].

[4] G. Sylvest, J. J. Andersen, K. V. Pedersen, M. H. Brun and P. Fremantle, "PEPPOL Transport Infrastructure: Lightweight Message Exchange (LIME)," 1 October 2010. [Online]. Available: https://joinup.ec.europa.eu/svn/peppol/PEPPOL_EIA/1-ICT_Architecture/1-ICT-Transport_Infrastructure/13-ICT-Models/ICT-Transport-LIME_Specification-101.pdf. [Accessed 4 July 2012].

[5] G. Sylvest, J. J. Andersen, K. V. Pedersen, M. H. Brun and P. Fremantle, "PEPPOL Transport Infrastructure: Secure Trusted Asynchronous Reliable Transport (START)," 1 October 2010. [Online]. Available: https://joinup.ec.europa.eu/svn/peppol/PEPPOL_EIA/1-ICT_Architecture/1-ICT-Transport_Infrastructure/13-ICT-Models/ICT-Transport-START_Service_Specification-101.pdf. [Accessed 4 July 2012].

[6] H. Lockhart, B. Campbell, N. Ragouzis, J. Hughes, R. Philpott, E. Maler, P. Madsen and T. Scavo, "Security Assertion Markup Language (SAML) V2.0 Technical Overview," 25 March 2008. [Online]. Available: https://www.oasis-open.org/committees/download.php/27819/sstc-saml-tech-overview-2.0-cd-02.pdf. [Accessed 4 July 2012].

[7] e-Freight, "About e-Freight," [Online]. Available: http://www.efreightproject.eu/default.aspx?articleID=18749&heading=The%20Project. [Accessed 4 July 2012].

[8] e-Freight, "Background," [Online]. Available: http://www.efreightproject.eu/default.aspx?articleID=18776&heading=Background. [Accessed 4 July 2012].

[9] M. Enered, S. Apel and E. Valtersson, "SITS Konceptbeskrivning WP4," 2011.

[10] PEPPOL, "About PEPPOL," [Online]. Available: http://www.peppol.eu/about_peppol. [Accessed 4 July 2012].

[11] C. Ciciriello, "PEPPOL Transport Infrastructure: Technical Overview," 2011. [Online]. Available: http://www.peppol.eu/about_peppol/promotional-material/20111122-peppol-transport-infrastructure-v1.17.pdf. [Accessed 4 July 2012].

[12] R. Steffen, J. Preißinger, T. Schöllermann, A. Müller and I. Schnabel, "Near Field Communication (NFC) in an Automotive Environment," 20 April 2010. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5476471. [Accessed 4 July 2012].

[13] TrackIT Systems, "Active and Passive RFID: Two Distinct, But Complementary, Technologies for Real-Time Supply Chain Visibility," [Online]. Available: http://www.thetrackit.com/library/Active%20vs%20PassiveRFIDWhitePaper.pdf. [Accessed 4 July 2012].

[14] J. Langer, C. Kantner and J. Scharinger, "NFC Devices: Security and Privacy," 7 March 2008. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4529403. [Accessed 4 July 2012].

[15] Wireless Technology Advisor, "RFID Range: And the Things That Affect It," [Online]. Available: http://www.wireless-technology-advisor.com/rfid-range.html. [Accessed 4 July 2012].

[16] M. Enered, "SITS Kommunikationsramverk WP4," 2011.

[17] L. R, "Near Field Communication in Business: How BlackBerry is changing the Landscape of Mobile Interaction," Blackberry, 20 June 2012. [Online]. Available: http://bizblog.blackberry.com/2012/06/nfc-blackberry-business-video/. [Accessed 4 July 2012].

[18] A. Mitrokotsa, M. Beye and P. Peris-Lopez, "Classification of RFID Threats based on Security Principles," 2011. [Online]. Available: http://www.lightweightcryptography.com/papers/Books/Mitrokotsa_SV2011.pdf. [Accessed 4 July 2012].

[19] "RFID-Zapper," 19 April 2007. [Online]. Available: http://events.ccc.de/congress/2005/static/r/f/i/RFID-Zapper%28EN%29_77f3.html. [Accessed 4 July 2012].

[20] J. I. Library, "Application of Attack Potential to Smartcards," 30 April 2006. [Online]. Available: http://www.ssi.gouv.fr/site_documents/JIL/JIL-The_application_of_attack_potential_to_smartcards_V2-1.pdf. [Accessed 4 July 2012].

[21] S. Kladko, "SPA and DPA: Possible Testing Solutions and Associated Costs," 21 September 2007. [Online]. Available: http://csrc.nist.gov/groups/STM/cmvp/documents/fips140-3/physec/papers/physecpaper08.pdf. [Accessed 4 July 2012].

[22] G. Sylvest, J. J. Andersen, K. V. Pedersen, M. H. Brun and M. Edwards, "PEPPOL Transport Infrastructure: Service Metadata Locator (SML)," 01 October 2010. [Online]. Available: https://joinup.ec.europa.eu/svn/peppol/PEPPOL_EIA/1-ICT_Architecture/1-ICT-Transport_Infrastructure/13-ICT-Models/ICT-Transport-SML_Service_Specification-101.pdf. [Accessed 4 July 2012].

[23] W. Stallings and L. Brown, Computer Security: Principles and Practice, Upper Saddle River, NJ: Pearson Prentice Hall, 2008.

[24] T. Cane, "Reference Solutions for Next Generation National Single Windows," 20 November 2011. [Online]. Available: http://www.efreightproject.eu/default.aspx?articleID=18896&heading=Deliverables. [Accessed 4 July 2012].

[25] Carnegie Mellon University, "CERT® Advisory CA-1998-01 Smurf IP Denial-of-Service Attacks," 13 March 2000. [Online]. Available: http://www.cert.org/advisories/CA-1998-01.html. [Accessed 4 July 2012].

[26] M. Kenney, "Ping of Death," 21 October 1996. [Online]. Available: http://insecure.org/sploits/ping-o-death.html. [Accessed 4 July 2012].

[27] G. Sylvest, J. J. Andersen, K. V. Pedersen, M. H. Brun and P. Fremantle, "PEPPOL Transport Infrastructure: Service Metadata Publishing (SMP)," 1 October 2010. [Online]. Available: https://joinup.ec.europa.eu/svn/peppol/PEPPOL_EIA/1-ICT_Architecture/1-ICT-Transport_Infrastructure/13-ICT-Models/ICT-Transport-SMP_Service_Specification-101.pdf. [Accessed 4 July 2012].

[28] C. Kolias, V. Kolias and G. Kambourakis, "A Secure and Efficient Authentication Protocol for Passive RFID Tags," 10 September 2009. [Online]. Available: http://www.icsd.aegean.gr/publication_files/conference/982366035.PDF. [Accessed 4 July 2012].

[29] S. Choi, S. Lee and H. Lee, "Security Enhanced Authentication Protocol for UHF Passive RFID System," 19 June 2011. [Online]. Available: www.thinkmind.org/download.php?articleid=icwmc_2011_14_30_20184. [Accessed 4 July 2012].

[30] Radboud University Nijmegen, "Security Flaw in Mifare Classic," [Online]. Available: http://www.sos.cs.ru.nl/applications/rfid/main.html. [Accessed 4 July 2012].

[31] Cisco, "Strategies to Protect Against Distributed Denial of Service (DDoS) Attacks," 22 April 2008. [Online]. Available: http://www.cisco.com/en/US/tech/tk59/technologies_white_paper09186a0080174a5b.shtml. [Accessed 4 July 2012].

[32] R. Mohan, "How to Defend Against DDoS Attacks," 27 April 2010. [Online]. Available: http://www.securityweek.com/content/how-defend-against-ddos-attacks. [Accessed 4 July 2012].

[33] ETSI, "Telecommunications and Internet converged Services and Protocols for Advanced Networking," March 2011. [Online]. Available: http://www.etsi.org/deliver/etsi_ts/102100_102199/10216501/04.02.03_60/ts_10216501v040203p.pdf. [Accessed 4 July 2012].

[34] G. Hancke and M. Kuhn, " An RFID Distance Bounding Protocol," 9 September 2005. [Online]. Available: http://www.cl.cam.ac.uk/~mgk25/sc2005-distance.pdf. [Accessed 4 July 2012].

[35] T. Loukusa and M. Simovits, “Så bräcklig är säkerheten i RFID-kretsar”, TechWorld, no. 9, pp.52-57, 30 October 2012.