NTNU
Norwegian University of
Science and Technology

SafeT-Next Generation Safety Assessment Framework for Railway: Development of a Framework for the Practical Implementation and Facilitation of STPA

Sunniva Benedicte Nygård Hansen

Master of Science in Mechanical Engineering
Submission date: June 2018
Supervisor: Mary Ann Lundteigen, MTP

Norwegian University of Science and Technology
Department of Mechanical and Industrial Engineering

RAMS
Reliability, Availability,
Maintainability, and Safety

SafeT- Next Generation Safety Assessment Framework for Railway

Development of a Framework for the Practical Implementation and Facilitation of STPA

Sunniva Benedicte Nygård Hansen

June 2018-06-22

MASTER THESIS

Department of Mechanical and Industrial Engineering
Norwegian University of Science and Technology

Supervisor: Mary Ann Lundteigen- NTNU
Co-supervisor 1: Terje Sivertsen- Bane NOR
Co-supervisor 2: Bjørn Axel Gran- IFE/NTNU
Co-supervisor 3: Øystein Skogvang- Safetec

Preface

This is a Master thesis in RAMS at NTNU as part of the study program Mechanical Engineering, carried out during the spring semester of 2018. The thesis is carried out in close collaboration with Bane NOR, Safetec, IFE, and NTNU, which all have contributed with theory on safety assessment methods and frameworks, STPA, RAMS and RAMS requirements. The case study considered in the project is developed by Bane NOR, while IFE has provided input on the STPA. There is no assumed background of the readers of this report, but it is recommended to have some knowledge about RAMS and safety assessment in general.

Trondheim, 2018-06-22

Sunniva Benedicte Nygård Hansen

Acknowledgment

I would like to thank the following persons for their great help during the writing process of the Master's thesis:

Mary Ann Lundteigen, main supervisor at NTNU
Terje Sivertsen, co-supervisor at Bane NOR
Bjørn Axel Gran, co-supervisor at IFE
Øystein Skogvang, co-supervisor at Safetec

S.B.N.H.

Remark

Given the opportunity here, the RAMS group would recognize Professor Emeritus Marvin Rausand for the work to prepare this template. Some minor modifications have been proposed by Professor Mary Ann Lundteigen, but these are minor compared to the contribution by Rausand.

Executive Summary

In this thesis, a systems-theoretic process analysis workshop has been planned and conducted in order to see if the method is advantageous to use at complex systems, and if the analysis reveals more hazardous scenarios than the traditional hazard identification methods. The results from the systems-theoretic process analysis workshop were used to create a framework for the practical implementation and facilitation of the method. In addition, the systems-theoretic process analysis-security methodology was studied in order to figure out how to include dangers associated with information and communication technology security in a systems-theoretic process analysis review.

The results obtained from the systems-theoretic process analysis workshop included 44 unsafe control actions that may put the system in a hazardous state where accidents occur. On the basis of these, the systems-theoretic process analysis identified 8 remaining unsafe control actions after considering safety barriers. The traditional hazard identification methods used were a hazard and operability study and a failure modes and effects analysis. The hazard and operability study identified 4 hazards, which were equivalent to 9 of the unsafe control actions identified in the systems-theoretic process analysis. The failure modes and effects analysis identified 8 failure modes, which were equivalent to 14 of the unsafe control actions identified in the systems-theoretic process analysis. Consequently, the systems-theoretic process analysis identified more unsafe control actions than the traditional hazard identification methods, and thus proved advantageous to use at complex systems.

The systems-theoretic process analysis framework was inspired by already existing frameworks used for hazard identification, and based on input from experts, the systems-theoretic process analysis workshop, and literature research. The final framework displays inputs and outputs, as well as ten main steps that describe the approach for conducting a systems-theoretic process analysis workshop. The steps in the framework are divided into a planning or preparation phase, an execution phase and a post work or follow-up phase.

The essential hazard identification tool used for the systems-theoretic process analysis was the control loop. The control loop, which lays the foundation for the whole analysis, fulfilled all the requirements to models given by Bane NOR, and therefore proved to be a suitable model structure for hazard identification.

Contents

Chapter 1

1. Introduction

1.1 Background

It is stated in the report "Safety- Requirements to models" that a safety assessment framework is necessary in the SafeT project to ensure efficient, reliable, safe and environment friendly transportation (Sivertsen, 2017c). Due to the potentially catastrophic consequences of functional failure in infrastructure systems, it is essential to demonstrate that the systems can safely be taken into use and that the required safety level can be maintained throughout their lifetime. In the report, it is also mentioned that the use of models is a central part of the SafeT approach, so therefore the selection, combination, adaption and further development of adequate types of models is a key success factor for the project (Sivertsen, 2017c). In order to minimize risks and to avoid potential accidents and hazards, suitable models must be implemented (Sivertsen, 2017c).

Bane NOR has recently decided to renew the signal systems on the railway. The goal is to replace today's outdated technology with the computer based system European Railway Traffic Management System (ERTMS). The ERTMS is a signalling system that is common to all European countries, and the implementation of the new system will lead to a more stable railway with increased safety, an increase in punctuality, and more capacity long term (Jernbaneverket, 2015). The implementation of the ERTMS will cause major changes to the train operation. One of these changes is that the old system used for blocking a work area by short circuit of track fields no longer can be used, and a new system is therefore required (Sivertsen, 2017a; 2017b). Consequently, a new solution for securing a work area has been introduced, and a safety assessment framework must be implemented to ensure that the new solution is safe (Sivertsen, 2017a). Several hazard identification methods have been used for the system in the SafeT project in order to identify all possible dangers related to the implementation of the new solution, but an STPA has not been tested for the system yet. Therefore, STPA is the method that will be studied in this thesis.

1.2 Problem Formulation

Systems-theoretic process analysis (STPA) is a method for hazard identification so new that not much practice on the implementation part of the method have been established yet, compared to other traditional hazard identification methods as Hazard and Operability Analysis (HAZOP), Failure Mode Effects Analysis (FMEA), Fault Tree Analysis (FTA), etc. (Leveson, 2013). As it is today, an STPA can only be conducted as a desktop analysis based on the theory provided on the topic. Within hazard identification, it is an established practice

to involve experts, which is done by gathering them in a form of workshop (Rausand, 2011). In order to establish the same practice for STPA, it is necessary to modify the theoretical STPA approach such that it can be used in a workshop context. There are few sources on this topic, and it is very unclear how STPA analyses that have been completed previously have handled this. Therefore, by planning, executing, and evaluating a workshop with experts in combination with literature research, a framework for practical implementation and facilitation of STPA will be developed. The result of the workshop will also be used to confirm or reject the claim by Leveson that STPA is advantageous to use at complex systems, and that an STPA reveals more hazardous scenarios than the other traditional hazard identification methods (Leveson, 2013).

In addition, the relationship between STPA and systems-theoretic process analysis for security (STPA-SEC) will be studied because the system for securing a work area includes information and communication technologies (ICT). STPA-SEC is an extension of STPA, and it identifies dangers in systems that are based on that type of technology. Therefore, in addition to developing a framework for practical implementation and facilitation of STPA, the STPA-SEC methodology will be studied in order to figure out how to include dangers associated with ICT security in an STPA review. Based on all of these findings, other possible purposes STPA can have in a development process of a new concept will be discussed. The Master's thesis is an extension of the project thesis, which focused on the theoretical foundation of STPA, and the focus of this thesis is on the practical implementation of STPA.

1.3 Related Work

1.3.1 Use of Models and Frameworks in the SafeT Project

There are already some risk assessments that have been performed in the SafeT project, as well as workshops focusing on hazard identification. The risk assessments that have been completed so far are HAZOP, FMEA, and the Systems Modelling Language (SysML) (Gran, Karpati and Hauge, 2018). In the HAZOP, the whole system for securing a work area was analyzed, while in the FMEA and the SysML only the "secure" function was analyzed.

1.3.2 What Remains to be Done?

There are still plenty hazard identification methods that have not been considered in the SafeT project yet for identifying relevant hazards in the new solution for securing a work area (Gran, Karpati and Hauge, 2018). STPA is one of the hazard identification methods that is relatively new, and has barely been tested in the railway industry in Norway. Therefore, a complete STPA is carried out in this thesis to be able to create a framework for the practical implementation and facilitation of the method for future use.

1.4 Objectives

The main objectives of this thesis are:

System Description

1. Describe the main system and the relevant functions. Identify main technical, human, and organizational elements involved in the relevant function. Illustrate how the system operates with suitable models.

Theoretical Basis and Gap Analysis

2. Explain the background behind STPA, STPA in relation to the case study, and the STPA and the STPA-SEC methodology.
3. Describe the phases in traditional risk assessment, and the safety assessment frameworks for HAZOP and FMEA.

Analysis Part

4. Plan and conduct an STPA in a workshop for the "secure" function in the system for securing a work area.

For the Results

5. Summarize results from the STPA workshop, and compare it with results from the HAZOP and the FMEA. Discuss similarities and differences based on the workshops that have been conducted.

6. Create a framework for practical implementation and facilitation of STPA, and discuss strengths and limitations related to the use of this framework. Identify topics for further investigation.

1.5 The SafeT Project

The description of the SafeT project is based on the reports provided by Terje Sivertsen about the SafeT project, as well as Bane NOR's requirements to models (Sivertsen, 2017b; 2017c).

1.5.1 Brief Description

It is expected that the infrastructure in the society is efficient, reliable, safe, and environment friendly. Therefore, when the society develops and adapts more technologies, the infrastructure must also implement advanced technical systems for supervision, protection

and control. Because of the catastrophic consequences of a failure in such systems, it is crucial to demonstrate safety in these systems throughout their lifetime.

The purpose of the SafeT project is to contribute to this, and it can be done by developing a framework for modelling of system design and risk. This framework use models that can be developed gradually, starting with system hazards and continuing with adding details that affect safety in the system. The choice of adequate models is therefore a key success factor for the project. This way all relevant safety requirements will be identified, and it is possible to fulfil the safety requirements, which again will lead to elimination and/or controlling of the hazards. Different ways to use the models must be integrated in a way that facilitates the overall safety demonstration and assessment.

1.5.2 SafeT Requirements to Models

The purpose of using models in the SafeT project is to be able to describe and analyze the structure and behaviour of a system, as well as the system's interaction with its environment. It is also used for supporting activities within risk assessment and hazard control, and to derive the necessary safety requirements to handle the hazards. In addition, an important purpose of the models is to communicate design and risk aspects, and the models are central in safety argumentation.

When finding the best suitable model for the project, there are many aspects that must be considered. The most important aspects are the requirements shown in Figure 1.1. These requirements are given to limit the scope to what are relevant and avoiding putting too much effort into irrelevant work. The full list of requirements is found in Appendix B.

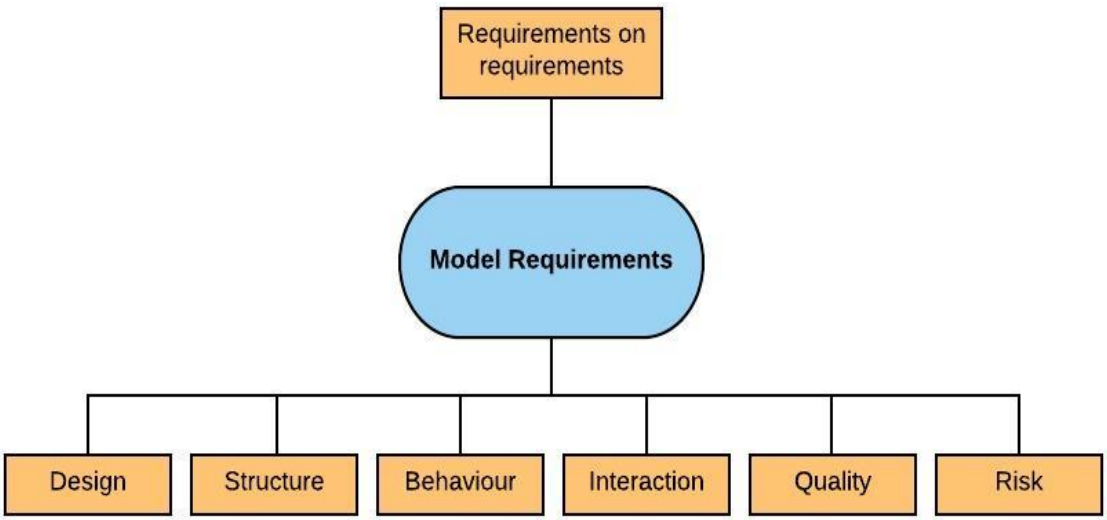

Figure 1.1: Bane NOR's requirements to models in the SafeT project

The set of requirements established for the SafeT project, concern different parts of the model. There are requirements on the structure and on the design of the model. There are also requirements on the behaviour and the interaction between the components in the model. In addition, there are requirements on the quality of the model, as well as risk requirements. There are even requirements on the requirements. Table 1.1 is adapted from "SafeT-Requirements to models", and it describes the purpose of the different categories of requirements (Sivertsen, 2017c).

Table 1.1: Description of the categories of requirements to models (Sivertsen, 2017c)

Requirement category	**Description**
Requirements	To identify and specify safety requirements
Design	To analyse the safety aspects of a design
Structure	To model the static aspects of a system at any system level
Behaviour	To describe the dynamic aspects of a system at any system level
Interaction	To describe the reciprocal impact between a system and its environment
Quality	To assure clarity, unambiguity, consistency, etc.
Risk	To carry out the risk assessment and hazard control

When the models have been selected, it is necessary to demonstrate that the models are successful and ready to use, which can be done by demonstrating that the models fulfil the requirements given. Therefore, it is safe to say that the requirements can be considered to be the evaluation criteria used for the selection of design and risk models. Sometimes multiple models are necessary in order to fulfil all the requirements, and it is not expected that one single model fulfil all the requirements alone.

1.5.3 Case Study

The case study given in this Master's thesis is to prepare an STPA workshop on the new solution of securing a work area, which include preparing the presentation used in the workshop and creating relevant models used for the analysis part of the STPA, as well as a framework that will be followed throughout workshop. After the workshop is completed, feedback from the participants will be collected together with own experiences in order to get a better foundation for creating a framework for practical implementation and facilitation of STPA.

A limited functionality in the system for securing a work area will be the study subject in the workshop, which is set to be the "secure" function. The reason for this is that it is a central function, which is linked to several of the other functions. The "secure" function should be analyzed in detail when conducting the STPA, and the results will be used to create a

framework for practical implementation and facilitation of STPA. The results and feedback from the workshop should make it possible to answer questions like when to involve experts, what preparations that need to be done, and how the workshop should be performed in order to find answers to key questions in the different steps. Furthermore, the framework should give a good pointer on advantages and disadvantages related to an STPA where experts are involved, and how to include dangers related to Information and Communication technology (ICT) security in an STPA. In addition, the experiences from the workshop should make it possible to identify what other purposes one can see that STPA may have in a development process of a new concept.

1.6 Research Approach

The research approach has consisted of both literature studies and an analysis of the "secure" function, as well as information from experts in the field with experience in hazard identification and STPA. A workshop was held early in the process in order to receive experts' opinions on the analysis of the system, and in order to have time to improve the framework suggested beforehand. Theory related to the practical implementation and facilitation of traditional hazard identification methods was studied, as well as the STPA approach and the STPA-SEC methodology.

Terje Sivertsen, the external supervisor from Bane NOR, had an important role in everything concerning the technical part of the system and the "secure" function. Bjørn Axel Gran, another external supervisor from Institute for Energy Technology (IFE), contributed to the STPA analysis, as well as providing input to the control structures and information on ICT security. Øystein Skogvang, the external supervisor from Safetec, provided general information about the planning and execution of workshops on hazard identification. All of the external supervisors were able to contribute with their experiences from previous workshops they had attended. The internal supervisor at NTNU, Mary Ann Lundteigen, played a central part in facilitating the workshop and contributed to the layout and implementation of the workshop, as well as the layout and design of the thesis. Figure 1.2 shows how the different supervisors contributed, as well as the key sources used.

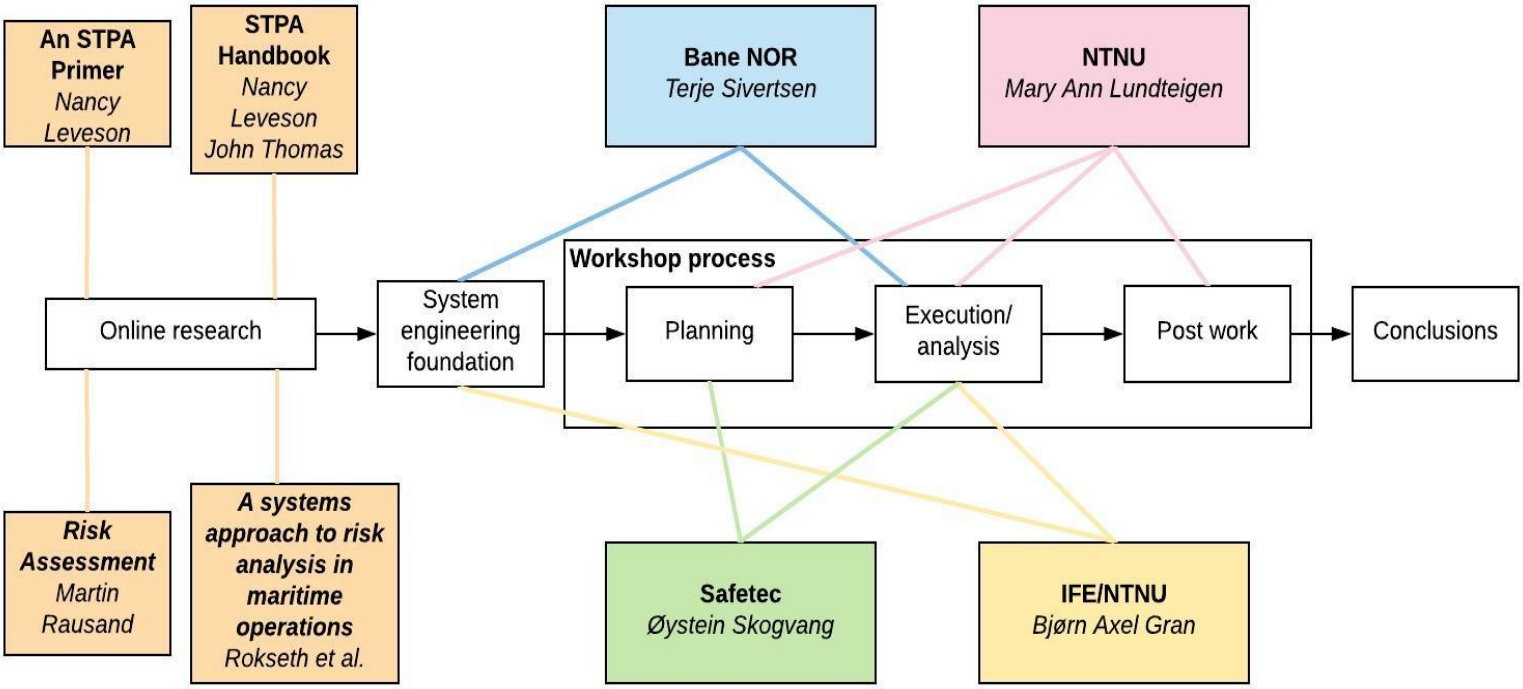

Figure 1.2: Contributions from the supervisors when writing the thesis

The main sources found on STPA were the reports published by Nancy Leveson, "An STPA Primer" (Leveson, 2013) and "STPA Handbook" (Leveson and Thomas, 2018). Relevant information on practical implementation of hazard identification methods was found in "Risk Assessment" by Martin Rausand (Rausand, 2011). Further on, "A systems approach to risk analysis of maritime operations" was used for creating the steps in the final framework (Rokseth et al., 2016), supplemented by the steps listed in the "STPA Handbook" (Leveson and Thomas, 2018). In addition, some information on the system and the STPA approach was adapted from the project thesis "SafeT- next generation safety assessment framework for railway", which described how to perform an STPA based on literature research only (Hansen, 2018).

1.7 Category of Hazard Identification Methods

STPA has several times proven to be theoretically more powerful through experimental application of STPA to different systems like the U.S. Missile Defence System and the Japanese spacecraft system. In all cases the STPA has identified all the hazardous scenarios and causes found in the traditional hazard identification methods, in addition to several new and undiscovered scenarios and causes (Leveson, 2013). When analyzing systems where safety is of big importance, the risk of overlooking hazards and thereby missing important safety requirements must be avoided (Sivertsen, 2017b).

An STPA can be conducted by one person only as well as by a group of people in a workshop. Since there are no official requirements concerning the participants' knowledge of the system or the STPA approach when conducting an STPA anyone can carry out such a workshop (Leveson, 2013). An STPA can be performed in weeks in the early phase of a project if experts in the field are involved, and the method is cost-effective, which makes STPA a good candidate for hazard identification on modern systems (Leveson and Thomas, 2018). Because STPA is a relatively new hazard identification method that is easy to implement (Leveson, 2013), and there is no framework for practical implementation that exists yet, it is of high interest to develop such a framework.

1.8 Limitations

There are limitations to the study and research approach regarding how the workshop is conducted, and the thoroughness of the analysis process. The results of the workshop are dependent on the knowledge and experience of the participants, the preparations done beforehand, and the models used when performing the analysis (Rausand, 2011). The fact that the participants were all familiar with the system and the "secure" function, as well as they had participated in workshops before, might have been an advantage and affected the result positively. On the other hand, having too much knowledge about the system and having participated in other workshops concerning the same system might make it difficult to discover new potentially dangerous scenarios. The models used for the analysis were made by the student, as well as all the post work, but with some input from experts. This is also a limitation to the study because it can make parts of the study subjective, and human analysis and human review is a source to inaccuracy (Leveson, 2013).

The fact that only a part of the system for securing a work area, the "secure" function, was studied is a limitation as well because other functions that are a part of the main system might affect the "secure" function as well. This might result in inaccuracy in the comparison of the STPA and the HAZOP and FMEA reports. Because there is no formal mathematical model for the entire system and for how it will operate, there is also a possibility of incompleteness in the study (Leveson, 2013).

The limitations of the study can be handled by reducing potential incompleteness and insecurities by structuring the process to optimize human review and processing (Leveson, 2013). The planning of the workshop is essential to handle limitations, and by having continuously reviewing and input from experts, the analysis can be optimized.

1.9 Outline

- Chapter 2. System description and system operation: This chapter describes the overall system used in the case study, and the different components it exists of. The "secure" function is described in detail, as well as the tasks and operators that are related to this function.

- Chapter 3. Theoretical background: This chapter gives the theoretical background needed on STPA in this report. It explains the theory behind STPA and STPA-SEC, as well as the different approaches. Furthermore, theory on facilitating a workshop based on traditional hazard identification methods is provided.

- Chapter 4. Suggested framework and the analysis: This chapter presents the suggested framework used in the workshop. It also includes the STPA analysis conducted in the workshop step-by-step, and it explains how guide words, figures, control loops, and an observation form were used to perform the analysis.

- Chapter 5. Result- comparison: In this chapter, the results from the STPA workshop are compared with the results from the HAZOP workshop and the FMEA workshop. Both similarities and differences are discussed.

- Chapter 6: Result- final framework: This chapter presents the final framework for practical implementation and facilitation of STPA divided into a planning phase, an execution phase, and a post work/evaluation phase.

- Chapter 7. Conclusion: This chapter includes conclusions, discussion and recommendations for further work with STPA. It summarizes what was found in chapter 4-6, and it is discussed whether the results are useful. In addition, the chapter deals with what could have been done differently.

- Bibliography

- Appendix A: Acronyms
  Appendix B: Full list of requirements given by Bane NOR
  Appendix C: Additional tables used in the STPA analysis
  Appendix D: Survey- Evaluation of the STPA workshop

Chapter 2

2. System Description and System Operation

In this chapter, the main system and the system operation are explained, as well as the components that are included in the system. The "secure" function, which is the study object in the analysis, is described in detail through tasks performed and operators involved in the function.

2.1 System Description

2.1.1 Work Area

The case study for this Master's thesis is securing a work area when maintenance work is carried out on the railroad tracks. Bane NOR defines a work area as "a track section (possibly more than one track) that can be disposed for work, without any trains entering or leaving the area." (Sivertsen, 2017a). Figure 2.1 is adapted from the report "Case example on securing work area" and shows how the work areas are divided in a crossing place on a railroad track.

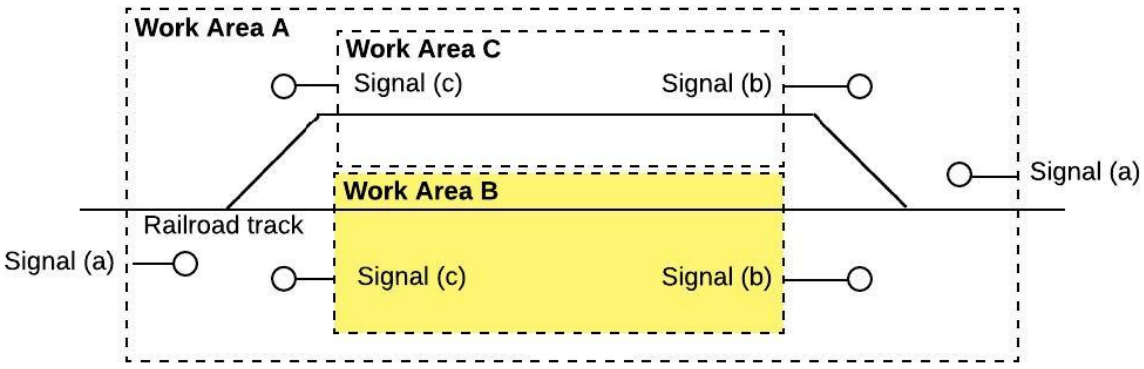

Figure 2.1: Work areas on a crossing place on the railroad track

At a work area, several persons are involved through dialogue, coordinated action, and mutual control, in addition to technical barriers designed into the system (Sivertsen, 2017a). Actors that are involved at a work area and their responsibilities are described in Table 2.1.

Table 2.1: Actors and their responsibilities

Actor	Description
Main Safety Guard (MSG)	Has the main responsibility of all the work areas, and communicates with the local safety guard
Local Safety Guard (LSG)	Has the responsibility of the current work area, and is responsible for the safety of the workers

Worker (W)	Maintenance workers perform their tasks on the railway tracks, and communicate with the LSG

2.1.2 New Solution for Securing a Work Area

The current solution used in Norway for securing a work area when axle counters are used involves removing a physical key for the relevant work area. The key can be removed from its lock when the train dispatcher has both blocked the work area and released the key. This solution is both expensive and inefficient due to the need of physical equipment along the tracks, as well as physically interlocking this with the signalling system (Sivertsen 2017b; 2017c).

In addition to the already existing automated protection systems like points, derailers, main signals, etc. (Sivertsen, 2017a; 2017b), it is suggested to implement a software system at the work area that will simplify the tasks and improve the safety on the work area (Sivertsen, 2017b). Therefore, a new solution for securing a work area is introduced, where the main principle is to reduce the physical measures needed in the infrastructure (Sivertsen, 2017a). This is done by marking the work areas with quick response (QR) codes, which is shown through work area B in Figure 2.2. The figure shows all the objects and actors involved at a work area.

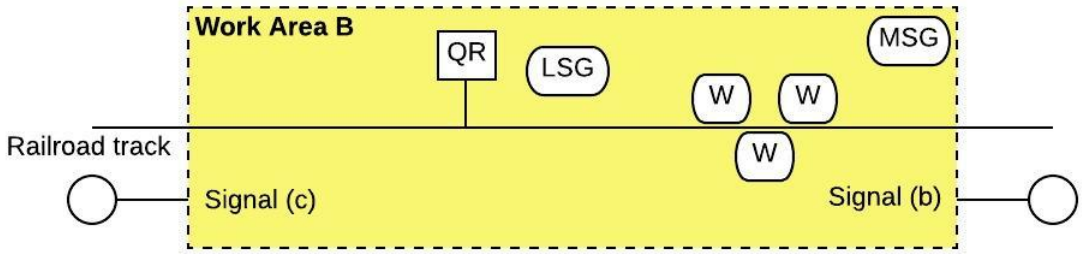

Figure 2.2: Objects and actors involved at a work area

The new system has several interfaces to its environment, and it includes technical systems, different categories of human operators, and the railway duty holder's organization (Sivertsen 2017a; 2017b). Figure 2.3 is adapted from the report "Case example on securing work areas", and it shows the roles and interfaces of the new solution introduced by Bane NOR (Sivertsen, 2017b). The interfaces between the operational support staff and the other roles are not included to simplify the figure.

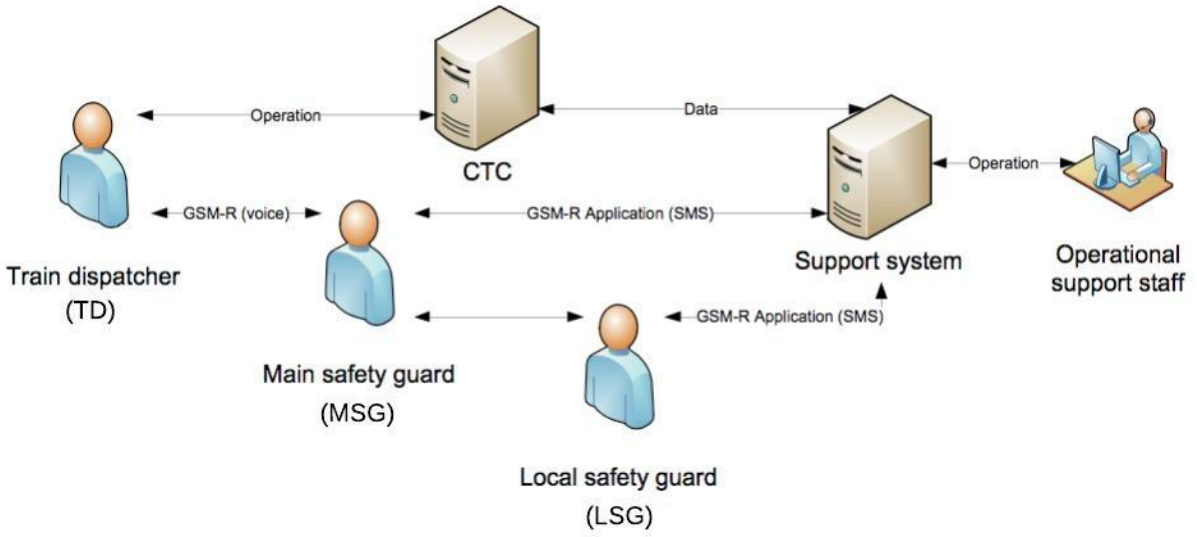

Figure 2.3: Roles and interfaces of the new solution (Sivertsen, 2017b)

Overall, the system shows how the human operators communicate with the centralized traffic control (CTC) system, the support system and the central computer through their smartphones, which all have Global System for Mobile Communications-Railway (GSM-R) receivers and transmitters. The GSM-R system contributes to digitalizing the railway, reduces operating cost, improves safety, and gives faster and more effective responses to hazards. In addition, it eliminates the need for drivers to exit the train if a problem occurs (Guide to GSM-R System, 2017). As seen from the figure, the main safety guard and the local safety guard communicates through SMS, while it is voice communication between the train dispatcher and the main safety guard. Data are transferred between the computers, while an operation is occurring between the operational support staff and the operational support computer.

2.2 System Operation

2.2.1 Main Functions

Bane NOR has specified twelve main functions for the new solution for securing a work area, which are applicable to the smartphone, and are listed in Table 2.2 in a random order (Sivertsen, 2017b). The “secure” function, which will be the study object in this thesis, is marked with green in the table.

Table 2.2: Main functions in the new solution for securing a work area (Sivertsen, 2017b)

Number	Function	Description

1	Log in	Logging into the system, thereby getting access to the other main functions.
2	Log out	Logging out of the system, thereby being prevented from using other functions before a new login.
3	Join	Enrolling in a work area, thereby preventing the safety guard in charge to release the work area
4	Resign	Withdrawing from a work area, thereby allowing the safety guard in charge to release the securing of the work area
5	Secure	Securing a work area, thereby preventing the work area from being unblocked
6	Release	Releasing a secured work area, thereby allowing the work area to be unblocked
7	Set time	Setting the time available for work in a work area, thereby allowing an automatic countdown of the time available
8	Time	Reading the time available for work in a work area, thereby facilitating management of work in the work area
9	Status	Reading the status of a work area, thereby facilitating management of work in the work area
10	Takeover	Requesting takeover of responsibility for a work area
11	Full takeover	Requesting takeover of another safety guard's responsibilities
12	Overview	Overview of the work areas the safety guard is in charge of or enrolled in

In addition, Figure 2.4 illustrates the main functions on the application in the correct order that they are performed, as well as which actor performs the different functions. The figure is adapted from Mary Ann Lundteigen, with some minor changes applied to it.

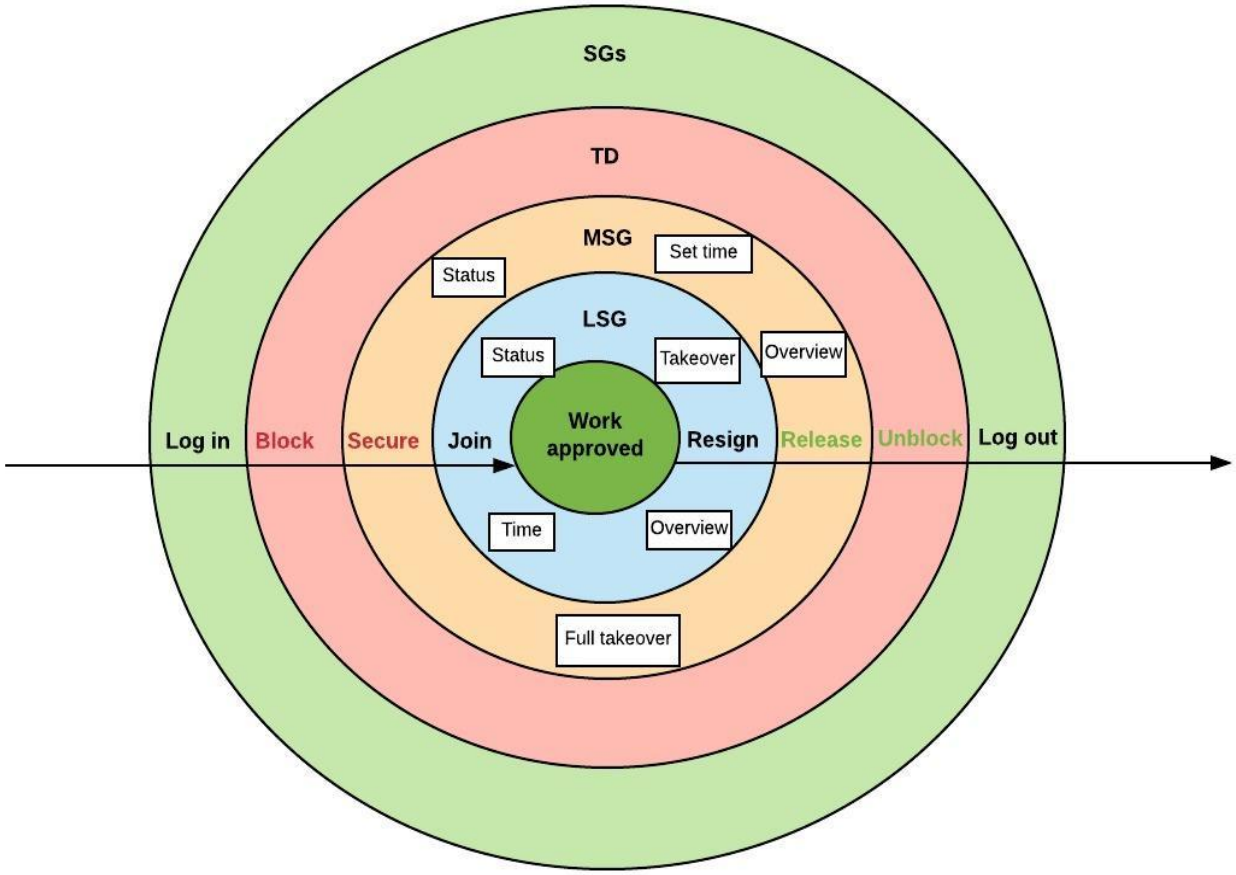

Figure 2.4: Main functions in the order they are performed (adapted from Mary Ann Lundteigen, 2017)

As seen in the figure, full takeover and set time are functions that are performed by the main safety guard, while takeover and time are functions that are performed by the local safety guard. Status and overview are functions that can be performed by all safety guards.

2.2.2 The "Secure" Function

As a part of the Master's thesis, an STPA workshop was prepared and executed. For this workshop, the "secure" function in the GSM-R system was chosen as the subject of the analysis. The reason this function was chosen is because it is a central function in the system, and if the securing is not performed, the whole procedure will stop. In addition, the securing can cause potentially dangerous scenarios, and is therefore an essential function for the hazard identification.

From Figure 2.4 it can be seen that it is the main safety guard only that can perform the two functions related to securing a work area, which are "secure" and "release". It can also be seen that in order for the main safety guard to be able to secure a work area, a train dispatcher

must block the work area first. The same applies to the "unblock" function. The work area must be released by the main safety guard before the train dispatcher can perform the function "unblock", and thus open the work area for incoming trains (Sivertsen, 2017b). Therefore, by securing a work area, the work area is also prevented from being unblocked. Bane NOR has divided the "secure" function into fifteen steps, which describe all the actions performed during securing. The fifteen steps are listed in Table 2.3, and also in Figure 2.5 that illustrates how the actors communicate, and the responsibilities of each actor (Sivertsen 2017b).

Table 2.3: The fifteen steps performed in the "secure" function (Sivertsen, 2017b)

Order	**Description**
1	The safety guard calls the train dispatcher and keeps the line until the work area is secured (using the communication as a barrier against hazards caused by human or technical failures).
2	The safety guard selects Secure from the application's main menu.
3	The application asks the safety guard to scan the work area code.
4	The safety guard scans the work area code.
5	The application sends a message to the support system, with a request to secure the work area and put the safety guard in charge.
6	The support system checks if the safety guard can secure the work area, which requires that the work area is not already secured. If the safety guard cannot secure the work area, the support system sends a message back to the application, which informs the safety guard that the securing request is rejected, and explains why.
7	If the safety guard can secure the work area, the safety system adds an entry associating the safety guard to the work area.
8	The train dispatcher blocks the work area, and releases it for securing.
9	The support system checks that the work area is blocked and released for securing. If the work area is not blocked and released for securing, the support system sends a message back to the application, which informs the safety guard that work area cannot be secured, and explains why.
10	The support system sends a confirmation to the application that the work area is secured, with a request that the safety guard confirms the work area.
11	The application asks the safety guard to scan the work area code.
12	The safety guard scans the work area code.
13	The application sends a message to the support system, with the confirmed work area.
14	The support system checks that the confirmed work area is identical to the requested work area. If the work areas are not identical, the support system sends

	an alarm to the application and the CTC that the safety guard attempts to secure wrong work area.
15	If the work areas are identical, the support system ensures that the work area cannot be unblocked from the CTC before the securing has been released by the safety guard in charge of the work area, sends a confirmation back to the application, which informs the safety guard that the given work area has been secured.

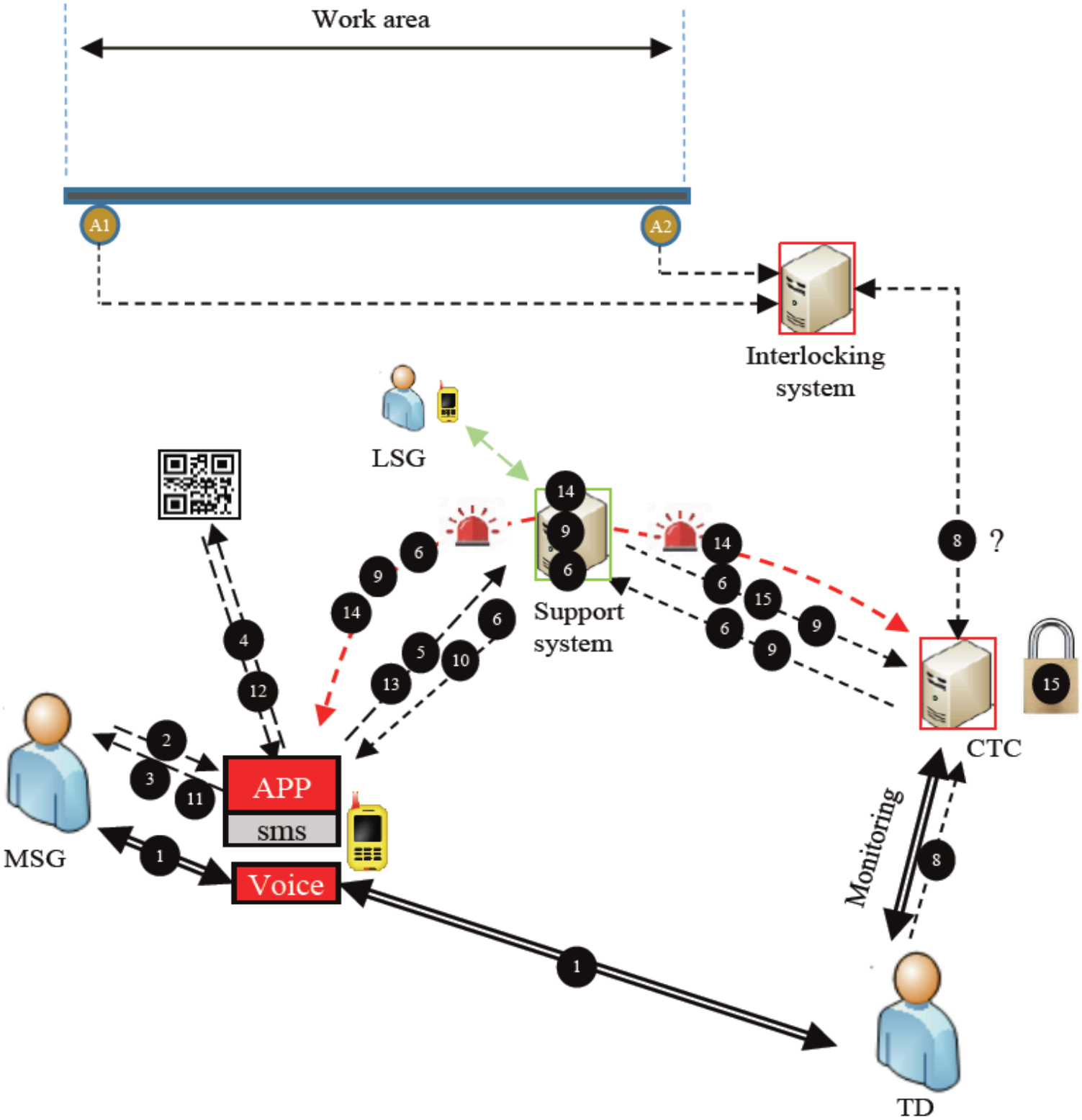

Figure 2.5: Steps in the "secure" function, and the actors involved in each step (adapted from Mary Ann Lundteigen, 2017)

2.2.3 Main Elements Involved in the "Secure" Function

In this section, a brief description of the roles of each of the component in the "secure" function is listed in order to provide an understanding of how this limited functionality of the GSM-R system operates. The list of components is based on Figure 2.5. The main elements are comparable to barriers, which are measures that will have as a function to protect when errors occur or in dangerous scenarios, and the main elements are divided into technical, human, and organizational elements (Eltervåg et al., 2017).

Technical

The technical elements are automated controllers involved in the "secure" function, and often equipment and systems that are included in the realization of a barrier function (Eltervåg et al., 2017). The descriptions of the technical elements for the "secure" function can be found in "Case example on securing work areas", and they are described as (Sivertsen, 2017b):

- Support System (SuS)
  Equivalent to the central computer in Bane NOR's system description, and it includes the operational support computer and the operational support staff. The support system communicates with the applications via a GSM-R receiver and transmitter, and the operational support computer is used by the operational support staff to operate the system.
- Centralized Traffic Control (CTC)
  Ensures correct interaction between the support system and the CTC system.
- Interlocking System (IS)
  An arrangement of signal apparatus that prevents conflicting movements through an arrangement of tracks such as crossings or junctions.
- Application (App)
  The application on the smartphones used by the human operators, which contains all the relevant functions used when securing a work area.
- QR code (QR)
  The identification code for the work area, which is scanned by the LSG. SG includes both MSG and LSG.

Human

Human elements or operators are personnel with defined roles or functions and specific skills (Eltervåg et al., 2017). The human elements are often used as extra safety barriers since they can check the work area physically. The train dispatcher has the possibility to block the work area, but the workers should be able to prevent the train dispatcher from unblocking the work area before the work area is finished if needed (Sivertsen, 2017a; 2017b). The operational support staff is not included as an own component in the "secure" figure since they do not

play a central role in the "secure" function, but they are included as a part of the support system instead. The human elements involved in the "secure" function are (Sivertsen, 2017b):

- Main Safety Guard (MSG):
  Has the overall responsibility of the work area, and uses the application on the smartphone to secure and release the work area depending on if maintenance work is performed there or not.
- Local Safety Guard (LSG):
  Enrols in work areas, uses the application on the smartphone to join work areas where work is needed, and to resign from work areas when the work is done.
- Train Dispatcher (TD):
  Blocks the work area when maintenance work is performed, and unblocks the work area when the maintenance work is completed.
- Maintenance workers (W):
  The maintenance workers along the railway track are influenced by the system, but not directly interacting with it, which is why they are not included in the figure. They indicate their position to the TD when performing maintenance work on the railway tracks, who can block the section to prevent trains from entering.

Organizational

The organizational elements involved in the "secure" function deal with leadership, decision making and structure, people, and work processes and systems (Eltervåg et al., 2017). The organizational elements involved in the "secure" function are (Sivertsen, 2017b):

- Requirements to specific skills
  To meet the requirements to specific skills, training of the safety guards and the workers included in the "secure" function are done, and performance measures are performed regularly.
- Procedures
  Procedures need to be carried out to secure the work area. An example of such procedures are relevant safety courses for all employees involved.
- Organizational structure
  The roles in the "secure" function must be clear. The organizational structure defines how task allocation, coordination and supervision are done at a work area, as well as the accountabilities for decisions (Eltervåg et al., 2017).
- Organizational behaviour or culture
  The organizational culture must consist of shared values and cultural assumptions in order to provide the basis for an effective decision making (Leveson and Thomas, 2018).

Chapter 3

3. Theoretical Basis and Gap Analysis

Chapter 3 provides the theoretical basis needed to perform the STPA, as well as a gap analysis to decide what remains to be done. A theoretical description of the STPA approach, the STPA-SEC approach, and the traditional hazard identification approaches must be provided before conducting the STPA in a workshop, and to be able to create a framework for practical implementation and facilitation of STPA.

3.1 STPA and the STPA Methodology

3.1.1 Background

Most of the theory on STPA comes from Nancy Leveson's report on STPA released in 2013, "An STPA Primer", and Nancy Leveson and John Thomas' report released in 2018, "STPA Handbook". STPA is a relatively new hazard identification method, and it is based on the extended model of accident causation (Leveson and Thomas, 2018). The method was developed because the traditional hazard identification methods like Fault Tree Analysis (FTA), Hazard and Operability Analysis (HAZOP), Failure Modes and Effects Criticality Analysis (FMECA), etc. did not address newer, more complex software systems that use new technology and have more focus on interactions. STPA differs from traditional hazards analysis methods because it is based on systems theory rather than reliability theory, as well as it uses a top-down approach. Leveson describes systems theory as theory that deals with modern systems that consider the whole of the system to be more than the components separately (Leveson, 2013).

STPA has the same goals as the traditional hazard identification methods, which is to identify dangerous scenarios that can lead to hazards, and then be able to eliminate or control the hazards (Leveson and Thomas, 2018). The analysis involves studying how the components are connected, and how the components interact with each other. There have been several cases where STPA has proven to be more cost-effective and been performed over a shorter period of time than the traditional hazard identification methods (Leveson, 2013). Therefore, it is of high interest in this Master's thesis to assess whether STPA as a hazard identification method has been sufficiently developed to be used in a workshop context in the same way that HAZOP and FEMA are used today.

3.1.2 Systems-Theoretic Accident Model and Processes (STAMP)

Previously, traditional chain-of-failure-event causality models have been used, where accidents always are caused by a chain of failure events over time, and each of the events lead to cause the next event. Figure 3.1 is adapted from the "STPA Handbook", and it illustrates the traditional chain-of-failure-event causality thinking, as well as showing how each event is the direct result of the preceding event(s) (Leveson and Thomas, 2018).

Figure 3.1: Chain-of-failure-event causality (Leveson and Thomas, 2018)

STPA is based on the new model of accident causation, Systems-Theoretic Accident Model and Processes (STAMP), which is an extension of the current accident models and is based on systems theory. In the STAMP model, accidents are not chain-of-failure-events, but they involve complex and dynamic processes (Leveson and Thomas, 2018). Accidents occur when the system enters a hazardous state, which is a result of inadequate control and violated safety constraints. Therefore, in a STAMP model, safety must be treated as a dynamic control problem (Leveson, 2013). Figure 3.2 is adapted from "An STPA Primer", and illustrates the STAMP thinking, which expands the traditional model of causality beyond chain-of-failure-event to include unsafe interactions among system components (Leveson, 2013).

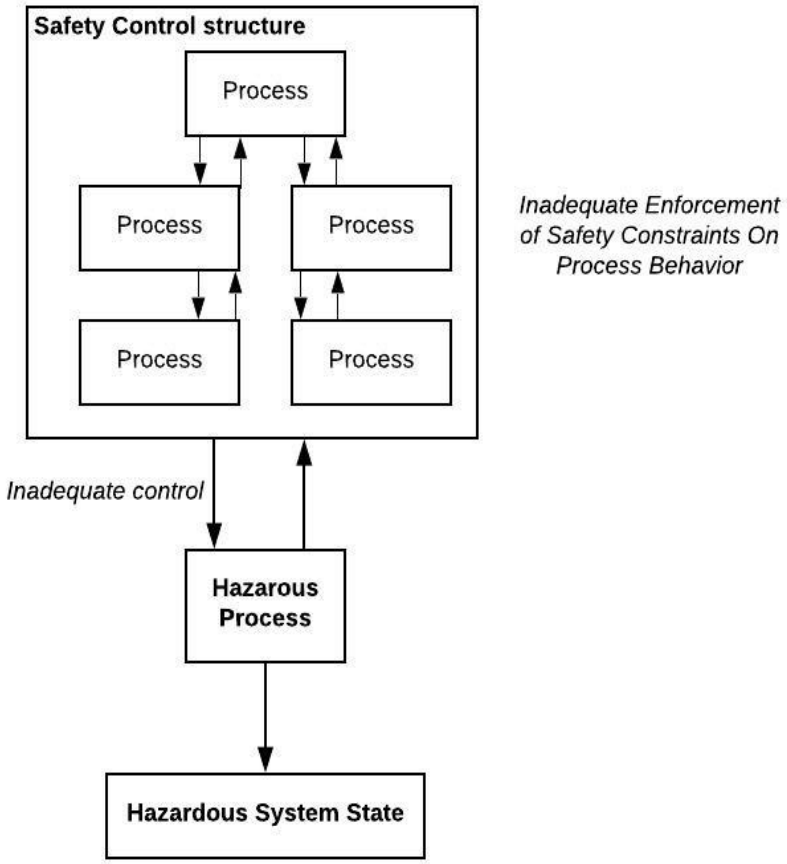

Figure 3.2: Systems-theoretic accident model and processes (STAMP) thinking (Leveson, 2013)

3.1.3 STPA in Relation to the Case Study

Leveson claims that there are plenty advantages of STPA over traditional hazard identification methods, which can be summarized in Table 3.1, which is also adapted from "STPA Handbook" (Leveson and Thomas, 2018).

Table 3.1: Description of the advantages of STPA (Leveson and Thomas, 2018)

Advantage	**Description**
Used on complex systems	"Unknown unknowns" that were previously only found in operations can be identified early in the development process, and either be eliminated or mitigated.
Can be started in early concept analysis	This means that it can also assist in identifying safety requirements and constraints, and design safety into the system architecture in early phase.
Includes software and human operators	The hazard analysis includes all potential causal factors in losses.
Provides documentation of system functionality	This type of documentation is often missing or difficult to find in large, complex systems.

Easily integrated	The STPA is easily integrated into the system engineering process and into model-based system engineering.

In previous evaluations and comparisons of STPA to more traditional hazard identification methods, STPA found all the causal scenarios identified in the traditional analyses in addition to many more (Leveson, 2013). Whether this allegation is true for the relevant system in this thesis will appear in the results when comparing the hazards identified in the STPA with the hazards identified in previous workshops on hazard identification.

In addition, the workshop completed as a part of the case study will demonstrate how an STPA works in practice as it is today. By using the experiences gained from the workshop and input from experts on the field together with the result of the STPA, a general framework for practical implementation and facilitation of STPA can be created.

3.1.4 STPA Approach

The theoretical STPA approach can be found by studying the STPA primer by Leveson (Leveson, 2013) and the handbook by Leveson and Thomas (Leveson and Thomas, 2018), as well as "A systems approach to risk analysis of maritime operations" (Rokseth et al, 2017). There are both similarities and differences concerning the content of the primer and the Handbook, which are listed in Table 3.2.

Table 3.2: Similarities and differences between the STPA primer and the STPA handbook

	An STPA primer	**STPA handbook**
Similarities • Content • STPA theory • Examples provided	• Hazards, control structures, UCAs, and safety requirements and constraints • Background, STAMP, and advantages/disadvantages • Plenty examples of how an STPA should be completed	
Differences • Year released • Structure • Focus • ICT security • Explanations	• 2013 (older) • More structured (step-by-step) • Theoretical use • Not included • Vague	• 2018 (newer, more updated) • Less structured • Theoretical/practical use • Included • Detailed

It has been decided to mainly follow the more detailed steps in the STPA approach suggested by Rokseth in "A systems approach to risk analysis of maritime operations" (Rokseth et al., 2017), supplemented by information from "An STPA primer" (Leveson, 2013) for the

analysis part of the workshop. The reasons for this are many. First of all, the STPA approach described by Rokseth and the approach in the primer is more structured than the approach described in the handbook by Leveson and Thomas, which is because the approach is listed as detailed steps and therefore makes it easier to follow. In addition, there are many examples of tables provided in the primer, which is useful when performing an STPA in practice. Nevertheless, since the handbook is newer and made for both theoretical and practical use of STPA, and contains some information on ICT security, it has been decided to include some main points and definitions from the handbook as well.

In addition to the six main steps adapted from "A systems approach to risk analysis of maritime operations", two extra steps concerning the post work were added, which are step 6 and step 8. The reason for this was to eliminate excessive unsafe control actions, as well as including the important evaluation process in a workshop context. An overview of the different steps and how they are performed can be found below (Rokseth et al., 2017).

Step 1: Describe the system and conceptualize it as a control system
In this step, the system is conceptualized as a control system, and the first version of the control loop is made. The way this is done sets the system boundaries and decides the scope for the analysis (Rokseth et al., 2017). In the STPA Handbook, a control loop is described as a system model that is composed of feedback control loops, and that an effective control structure will enforce constraints on the behaviour of the overall system. The control loop is pictured as a hierarchical control structure, and in general it consists of at least five types of elements, which are (Leveson and Thomas, 2018):

- Controllers
- Control Actions
- Feedback
- Other inputs to and outputs from components
- Controlled processes

The controllers provide control actions to control some processes in the system, and process models are then used to make decisions. The process models are the controllers' internal beliefs, and they may be about the actual process being controlled or about other aspects of the system or environment. The vertical placement of the controllers indicates the level of control and authority within the system. All downward arrows in blue represent commands or control actions, while all upward arrows in red represent feedback. These arrows will help manage complexity and recognize control relationships between different controllers (Leveson and Thomas, 2018). The generic model of a control loop is adapted from "STPA

Handbook" and modified to include all important details in a control loop. It is illustrated in Figure 3.3.

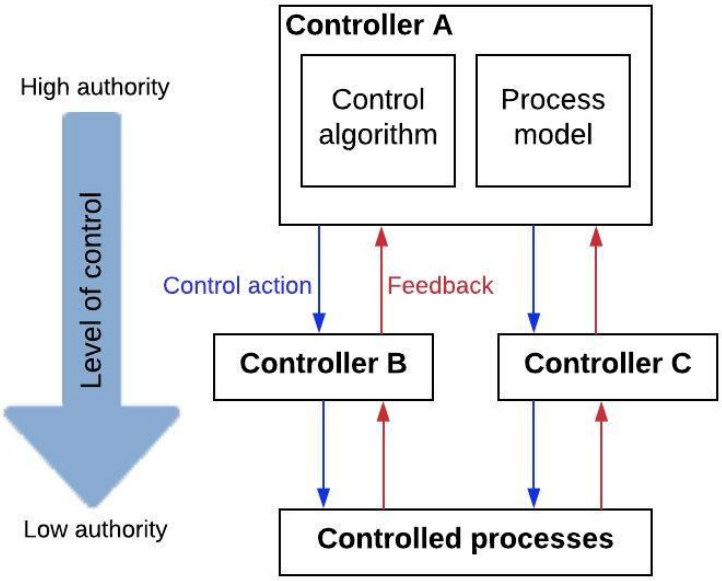

Figure 3.3: Generic model of a control loop (Leveson and Thomas, 2018)

By drawing a control loop, previously undiscovered flaws can be discovered immediately (Leveson and Thomas, 2018). The process of making a control loop might be a long process because the control loop is updated all the way during the analysis after new input or experiences are added.

Step 2: Identify System-Level Accidents (SLA), System-Level Hazards (SLH), and System-Level Safety Constraints (SLSC)

In step 2, the purpose and the goals of the system are identified by identifying the system-level accidents, hazards, and safety constraints. The definition of a system-level accident given in "An STPA Primer" is: "An undesired and unplanned event that results in a loss, including a human loss of human life or human injury, property damage, environmental pollution, mission loss, financial loss, etc." (Leveson, 2013). In the analysis performed in this thesis, it is decided that only loss of human life or human injury or damage to property or equipment are considered as system-level accidents.

A system-level hazard in an STPA is defined by Leveson as:" A system state or set of conditions that together with a worst-case set of environmental conditions, will lead to an accident (loss)." (Leveson, 2013). An important difference between traditional hazard identification methods and the STPA is the definition of hazards. Bane NOR classifies hazards as failures, errors or faults in the components (Sivertsen, 2014). In the STPA failures, errors or faults in the components are classified as causes of a hazardous state, not the state

itself. General words should be avoided in the STPA because they provide little information on the actual cause (Leveson, 2013).

Safety constraints are described by Leveson as constraints or requirements to the components in the system, and the enforcement of these safety constraints should keep the system away from a hazardous state. System-level safety constraints are described as requirements to the system on a higher level based on identified accidents and hazards only (Leveson, 2013).

Step 3: Identify controller responsibilities and process models

Once the controllers have been identified, responsibilities and process models can be assigned to each controller. Controller responsibilities are seen as high-level requirements for the controllers. The responsibilities are refinements of the safety constraints, and they tell us what each controller needs to do in order for the safety constraints to be enforced (Leveson, 2013). Next, control actions are formed and the feedback is derived from these responsibilities by identifying the process models that the controllers need to make decisions (Leveson and Thomas, 2018). Control actions are defined by Leveson as mechanisms that make changes to the inputs, while the process models are defined as possible inputs to the controller (Leveson, 2013). Flaws in the process models alone might result in unsafe control actions, and therefore studying the process models is a central part of this step.

Step 3 is of big importance to the focus of the analysis since it will influence the next step in terms of which control actions that will be analyzed further (Leveson, 2013).

Step 4: Identify potentially unsafe control actions (UCA)

The purpose of step 4 is to identify incidents where inadequate control can occur, which are called unsafe control actions (UCAs) (Rokseth et al., 2017). Leveson describes five generic modes of unsafe control actions (Leveson, 2013), which can be divided again to get more detailed modes of unsafe control actions. After dividing, there is a total of eight generic modes:

1. An action is not provided
2. An action is provided, but not followed
3. An unsafe action is provided
4. An action is provided too early
5. An action is provided too late
6. An action is provided in wrong sequence
7. An action is provided too long
8. An action is provided too short

By studying the responsibilities of each controller together with the eight possible modes of unsafe control, unsafe control actions for the system can be identified (Rokseth et al.2017).

Step 5: Identify scenarios and causal factors

In step 5, each part of the control loop is investigated in order to find how each of the UCAs could possibly occur. The process models are often involved in the dangerous scenarios, so it is of high importance to study the process models in the control structure (Leveson and Thomas, 2018).

Scenarios are manners in which the UCAs may occur, while causal factors are the reasons for why the scenarios may take place (Rokseth et al., 2017). Causal factors are what Bane NOR defines as hazards (Sivertsen, 2014).

Step 6: Identify remaining unsafe control actions

In order for the unsafe control actions identified in the STPA to be realistic, the remaining UCAs when considering both already existing safety barriers and planned safety barriers must be identified. This is an extra step that has been added to the original steps in "A systems approach" to improve the results and the reliability of the analysis (Rokseth et al., 2017).

Step 7: Identify Safety Constraints (SC)

The last step of the analysis is to develop safety constraints at the UCA level, scenario level, and safety constraints related to each causal factor. UCAs are used together with information on how and why they might occur in order to formulate safety constraints (Leveson, 2013). The safety constraints can later be used for tracing a particular UCA, which further leads to finding the system-level accidents (Leveson and Thomas, 2018).

The purpose of the safety constraints is to design strategies to avoid the UCAs (Rokseth et al., 2017). The safety constraints are defined as risk reducing measures by Bane NOR in other hazard identification methods (Sivertsen, 2014).

Step 8: Evaluation of the analysis

Step 8 is another step that has been added to the original STPA approach described in "A systems approach" (Rokseth et al., 2017). The reason for this is that the reflection on the results of the analysis is an important part when developing a framework for the practical implementation of STPA. An evaluation can be both qualitative and quantitative, and some examples of how the evaluation can be executed are by direct feedback, surveys, discussions, etc.

Further on, theory on practical implementation of STPA must be studied to be able to create a framework for the practical implementation of this method. Not much theory exists on this topic, but some steps on how to do a basic STPA are provided in "STPA Handbook" by Nancy Leveson and John Thomas. The basic steps when performing an STPA in practice are shown in Figure 3.4, complemented by an example drawing, which is adapted from the STPA handbook and modified as well (Leveson and Thomas, 2018).

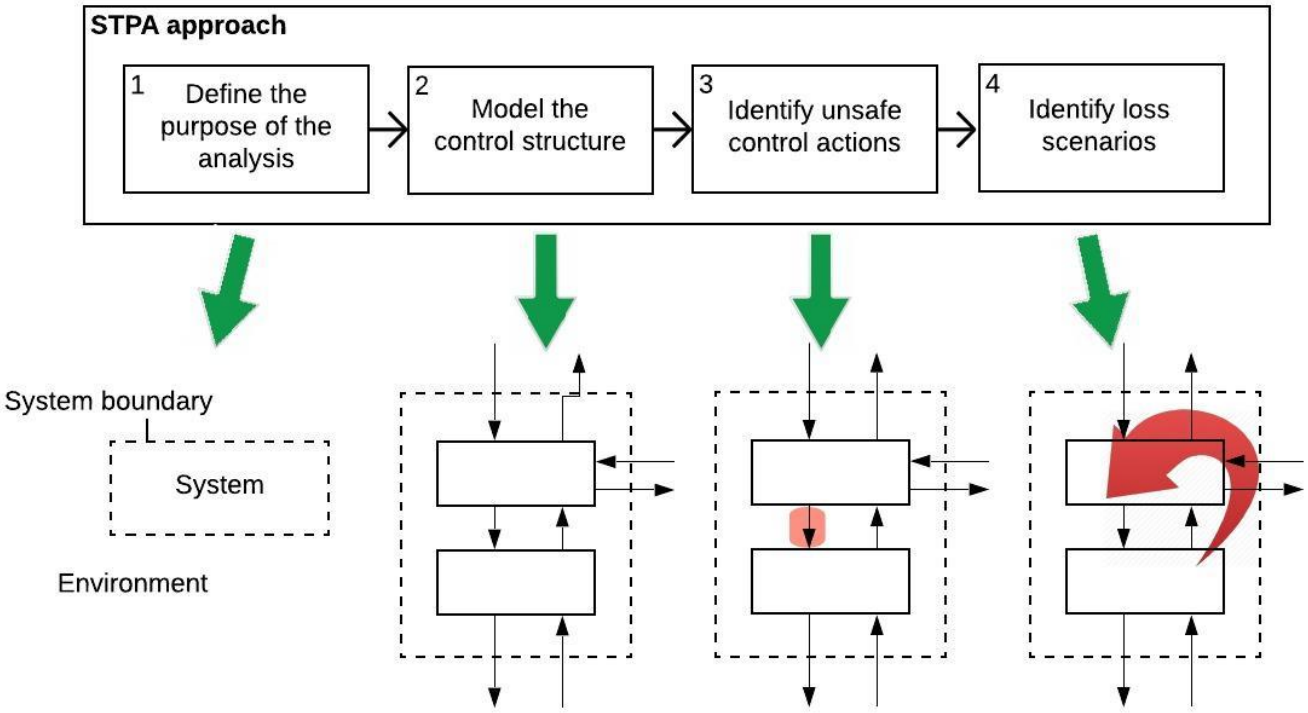

Figure 3.4: Basic steps in a practical STPA (Leveson and Thomas, 2018)

The basic steps for the practical implementation of STPA can be found below (Leveson and Thomas, 2018).

Step 1: Define the purpose of the analysis

The purpose of the analysis is the first step in any analysis method, and it includes the objectives of the analysis and the kind of losses the analysis aims to prevent. When considering this in the first step, fundamental questions that must be addressed are if only loss of human life should be considered or if it is important to include other aspects like security, reliability or other system properties. Defining the purpose of the analysis can be divided into four parts (Leveson and Thomas, 2018):

- Identify losses
- Identify system-level hazards
- Identify system-level safety constraints
- Refine hazards (optional)

Step 2: Model the control structure

In step 2, a model of the system must be made, which is called a control structure in an STPA. The purpose of the control structure is to show the relationships and interactions between the components in the system by modelling the system with feedback control loops. Usually, the control loop starts at a high level, and as new details are added to the system the control structure is updated and developed further. Nevertheless, the control structure does not have to be hierarchical. Modelling the control structure consists of the followings steps (Leveson and Thomas, 2018):

- Define control actions for each component
- Define process models for each component
- Define feedback used to observe the controlled process
- Model the control structure

Step 3: Identify unsafe control actions

The third step when performing a basic STPA analysis is to identify the unsafe control actions, which are found by studying the control actions and how they may lead to losses. Further on, the unsafe control actions are used to create safety requirements and constraints for the system. The steps in identifying unsafe control actions are (Leveson and Thomas, 2018):

- Consider each control action, and all the ways a control can be unsafe (eight generic modes of unsafe control are listed in the STPA approach)
- Identify process model flaws
- Define controller constraints

Step 4: Identify loss scenarios

In the fourth and last step, reasons for why unsafe control might occur in the system are identified, and possible dangerous scenarios are listed. Typical scenarios might be incorrect feedback, component failures, and inadequate requirements. The two main types of loss scenarios can be considered, and detailed scenarios related, are (Leveson and Thomas, 2018):

- Why would Unsafe Control Actions occur?
    - Failures related to the controller (for physical controllers)
    - Inadequate control algorithm
    - Unsafe control input
    - Inadequate process model
    - Feedback or information not received
    - Inadequate feedback is received

- Why would control actions be improperly executed or not executed, leading to hazards?
    - Control action not executed
    - Control action improperly executed

3.2 STPA-SEC

The theory presented on STPA-SEC is mainly based on "Systems thinking for safety and security" (Leveson and Young, 2013), and it explains key concepts in an STPA-SEC, as well as how to perform an STPA-SEC. Furthermore, "STPA-SAFESEC: Safety and security analysis for cyber-physical systems" (Friedberg et al., 2017), explains why the STPA-SEC was introduced, as well as describing some parts of the STPA-SEC approach. In addition, "STPA-SEC for cyber security/mission assurance" (Leveson and Young, 2014), and "System-Theoretic Process Analysis for Security (STPA-SEC): Cyber Security and STPA" (Young and Porada, 2017) cover some aspects of the STPA-SEC.

3.2.1 Background

Because today's critical infrastructures are becoming more interconnected and thus more complex cyber-physical systems (CPS), ensuring both their safety and security becomes of high importance (Leveson and Young, 2014). A cyber-physical system is defined by Friedberg as: "Physical processes and components that are connected over information and communication technologies (ICT), which are critical for correct system operation" (Friedberg et al., 2017). Traditional risk assessment methods do not handle the complexity of emerging CPS systems very well since they are based on the chain-of-failure-event causality models, which is why STPA-SEC was introduced to manage the failures related to cybersecurity in interconnections (Friedberg et al., 2017).

Cybersecurity threats are becoming a big concern to CPS systems, and there are several examples of how cyber-attacks have caused big damage. A cyber-attack is defined as an attempt to destroy or damage a computer network system, usually done by hackers (Leveson and Young, 2013). Within the railway sector there are still many obsolete technologies such as GSM-R, circuit-switching, and complete embedded systems (Sivertsen, 2014). Therefore, more modern technologies are adopted by actors in the railway industry in order to follow today's technological development. By adopting more modern technologies, the risk of cyber-attacks happening also increases (Friedberg et al., 2017). Because the system under consideration in this Master's thesis is a railway system that is adopting a modern technology, it is natural to address the security issues for that system as well. The security issue can be addressed by performing an STPA-SEC, which addresses the growing problem of securing

CPS systems against malicious attacks and international disruptions (Leveson and Young, 2013).

3.2.2 Key Concepts

Several aspects can be considered when defining the purpose of the STPA, and one of these are security. The STPA-SEC is an extension of STPA, and therefore the STPA-SEC approach is similar to the "ordinary" STPA approach with only a few changes made. It is therefore relevant to study the key concepts used for an STPA-SEC, and how the concepts differ between the STPA-SEC and the "ordinary" STPA. Table 3.3 shows the corresponding key concepts in the STPA and STPA-SEC (Leveson and Young, 2013). The definitions of the terms in an STPA is adapted from "An STPA Primer" (Leveson, 2013), while the terms used in an STPA-SEC is obtained from "Systems thinking for safety and security" (Leveson and Young, 2013).

Table 3.3: Key concepts in an STPA and an STPA-SEC

Key concept	**Meaning in STPA**	**Key concept**	**Meaning in STPA-SEC**
Security	A condition that results from the establishment and maintenance of protective measures	Cybersecurity	Prevention of damage to, and protection of systems against intentional disruptions
Loss	Loss in human life or injury, property damage, environmental pollution, mission loss, financial loss, etc.	Loss	Lack of control in computer systems
System accident	An undesired and unplanned event that results in a loss	System accident	Intentional digital disruptions/cyberattack
System hazard	A system state or set of conditions that together with a worst-case set of environmental conditions, will lead to a loss	System hazard	A situation that poses a level of threat to cyber-security.
Unsafe control action	Incidents where inadequate control can occur	Unsecure control action/vulnerable state	A weakness which can be exploited by a Threat Actor, such as an attacker, to perform unauthorized actions within a computer system.
Dangerous scenario	Manners in which the UCAs may occur	Disruption scenario	Manners in which the system is in a vulnerable state
Safety constraint	Risk reducing measures to avoid the UCAs	Security constraint	Requirements that prevent the disruption scenarios

			from happening (ex. access control)

3.2.3 The STPA-SEC Methodology

Both STPA and STPA-SEC has the STAMP accident causation model as a foundation, but STPA-SEC is an extension of the "ordinary" STPA hazard analysis (Young and Porada, 2017), which is illustrated in Figure 3.5.

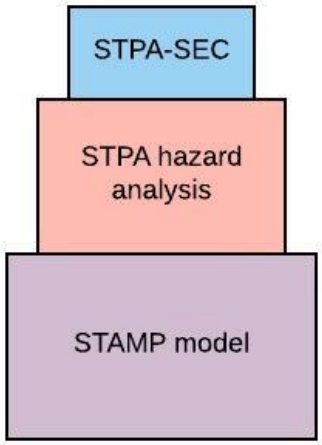

Figure 3.5: The relationship between STAMP, STPA, and STPA-SEC (Young and Porada, 2017)

The STPA-SEC can be performed both for abstract and physical systems with the goal to develop systems that enable us to more securely satisfy needs. It is a top-down approach, and therefore it can be applied to the beginning of the project (Leveson and Young, 2013). The STPA-SEC addresses both technical and organizational issues, and it identifies security vulnerabilities, security requirements, and scenarios leading to violation of security constraints (Friedberg et al., 2017).

In "System-Theoretic Process Analysis for Security (STPA-SEC): Cyber Security and STPA", the STPA-SEC approach is described through six main steps (Young and Porada, 2017), which are described in Table 3.4.

Table 3.4: Description of the main steps in an STPA-SEC (Young and Porada, 2017)

Step	Description
1	Define system purpose and goal • Define and frame security problem • "A system to do what, how, and why"
2	Identify unacceptable losses • Identify accidents/losses
2	Identify accidents and hazards • Identify system hazards/constraints
3	Create functional control structure • Identify model elements • Identify each model element's responsibilities

	• Identify control relationships • Identify control actions • Develop process model description • Identify process model variables • Identify process model variable values • Identify feedback providing PMV values • Check functional control structure model for completeness
4	Identify unsafe/unsecure control actions • Model functional control structure • Identify unsafe/unsecure control actions
5	Identify causal scenarios and causal factors • Trace hazardous control actions using information from life cycle • Identify scenarios leading to unsafe control actions • Identify scenarios leading to unsecure control actions • Place scenarios on a chart to ID more critical security scenarios
6	Mitigations and controls • Investigate security scenarios to select control strategy • Develop new security requirements, control, and design features to eliminate or mitigate unsafe/unsecure scenarios

The “ordinary” STPA and STPA-SEC share most of the basic steps, but the results and the procedures are different. In Figure 3.6, the STPA-SEC approach is compared to the “ordinary” STPA approach in order to detect the main differences between the two approaches. The “ordinary” STPA approach is adapted from “A systems approach to risk analysis in maritime operations” (Rokseth et al., 2017) and “An STPA Primer” (Leveson, 2013), while the STPA-SEC approach is adapted from “STPA Handbook” (Leveson and Thomas, 2018), as well as “Systems thinking for safety and security” (Leveson and Young, 2013).

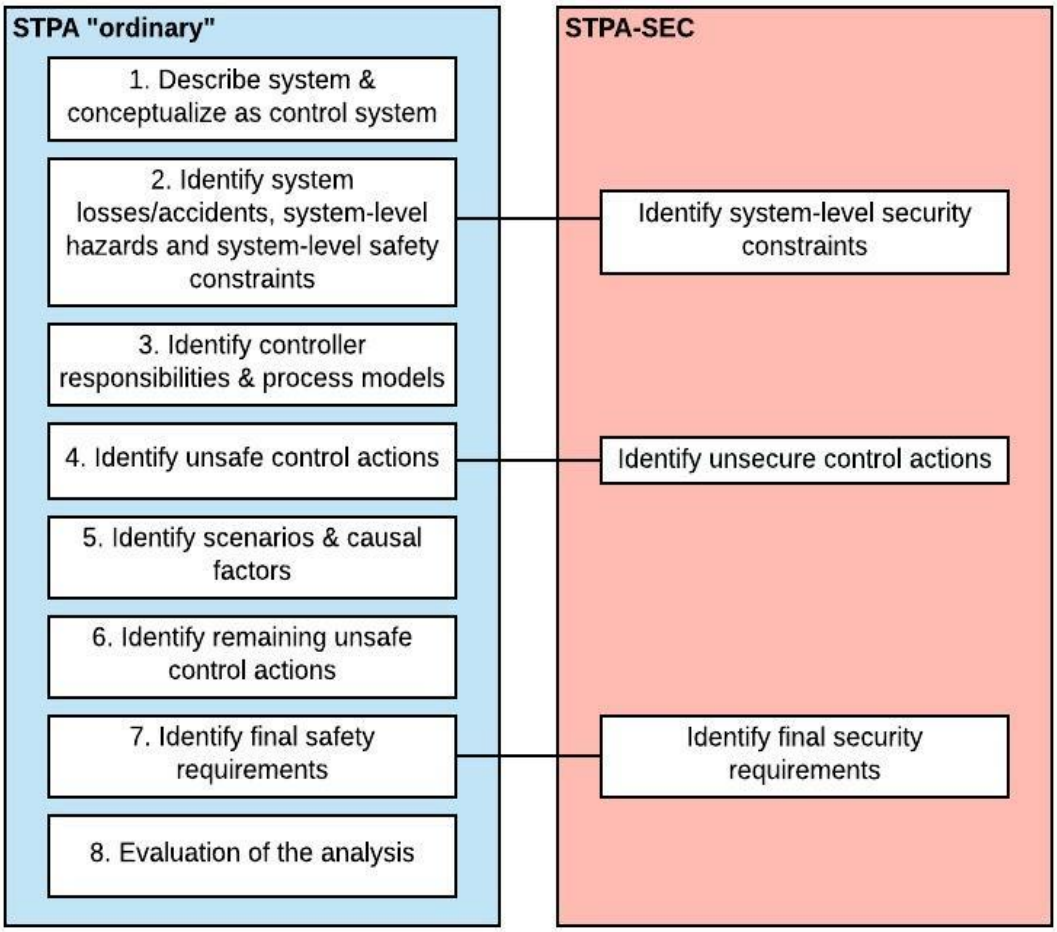

Figure 3.6: The "ordinary" STPA approach, and the STPA-SEC approach

As seen from Figure 3.6, some aspects must be added to the "ordinary" STPA approach when performing and STPA-SEC as it is an extension of the "ordinary" approach (Leveson and Thomas, 2018). In all of the steps, the focus is different when performing an STPA-SEC. The system will be described as a control system by including the parts of the system that are related to the security of the system. In addition, security will be the focus when identifying unacceptable losses and vulnerable states such that functions and system services can be located to be protected and controlled. The responsibilities and process models identified will also be related to security issues only (Leveson and Thomas, 2018). However, the "ordinary" STPA approach is extended to include the system-level security constraints, unsecure control actions, and the final security requirements when performing and STPA-SEC (Leveson and Young, 2013). The vulnerability lies in the interactions between the components, and it only appears in worst-case scenarios. The unsecure control actions are identified in the same way as in the "ordinary" STPA, which is by relating the different control actions to all modes of vulnerability (Leveson and Thomas, 2018). Further on, the unsecure control actions are used to create security requirements that will prevent disruption in the system (Leveson and Young, 2013).

STPA-SEC does not provide any answers to what measures that should be taken, which means that it is the security specialists' responsibility to identify reasonable protection mechanisms for the system (Leveson and Young, 2013). Nevertheless, STPA-SEC is today used for business and mission analysis, security control, hazard analysis, and especially on cybersecurity systems (Young and Porada, 2017).

3.3 Traditional Risk Assessment Frameworks

Most traditional risk assessment methods date from more than 50 years ago, which means that the frameworks and procedures used within these methods are well established today (Leveson, 2013). Therefore, it is necessary to study the existing frameworks and procedures in order to be able to develop a safety assessment framework for a newer hazard identification method.

Figure 3.7: Generic model for a safety assessment framework

Most of the safety assessment frameworks today are divided into a planning/preparation phase, an execution phase, and a post work or evaluation phase (Stangeland, Lundon and Skogvang, 2018). All of these phases are connected with smooth phase transitions over time, as illustrated in Figure 3.7. Because both a HAZOP analysis and an FMEA analysis have been conducted previously on the system under consideration, it will be useful to further study the approaches used in these methods. The theoretical foundation on HAZOP and FMEA is based on the book "Reliability, Maintainability and Risk" by David Smith (Smith, 2011), and the book "Risk Assessment" by Martin Rausand (Rausand, 2011).

3.3.1 Hazard and Operability Study (HAZOP)

A Hazard an Operability study (HAZOP) is a systematic hazard analysis process developed in the 1970s, and it was developed initially to be used during the design phase, but can also be applied to systems in operation (Rausand, 2011). In "Reliability, Maintainability, and Risk", a HAZOP is described as: "A study carried out by a multidisciplinary team, who apply guidewords to identify deviations from the design intent of a system and its procedures. The team attempt to identify the causes and consequences of these deviations and the protective

systems installed to minimize them and thus make recommendations which lead to risk reduction" (Smith, 2011).

Usually, a HAZOP team exists of a group of 5-8 people where ate least one member must have the authority to make decisions that may affect the system. The analysis part is done by a series of meetings as a guided brainstorming based on a set of guidewords, which will vary with the type of HAZOP conducted. The HAZOP leader is in charge of asking questions to stimulate the discussion by using guidewords. The purpose of the guidewords is to stimulate individual thought and to engage all the team members in the discussion (Rausand, 2011). These guidewords are applied to each of the process parameters or system modes in order to make it easier to discover potentially dangerous hazards in a team or a workshop. Each deviation of a parameter has a cause, and the causes lead to consequences, which must be assessed. Likelihood and severity may also be included in the HAZOP, depending on the scope of the analysis (Smith, 2011).

When performing a HAZOP, it is a requirement that all participants have a full knowledge of the operating system, which is why a HAZOP is usually conducted by a team that only consists of experts within the field. The HAZOP leader, who usually is the facilitator, should be independent of the project (Rausand, 2011). Nevertheless, the leader must be familiar with the system design, as well as having experience of HAZOP to be able to bring a vide view to the process. In addition, there should be a HAZOP secretary responsible for producing the record of the team's discussions and decisions (Smith, 2011).

The HAZOP approach is adapted from "Risk Assessment" by Rausand and divided into eight steps (Rausand, 2011), and modified by applying the three main phases for safety assessment frameworks (Stangeland, Lundon and Skogvang, 2018). The steps in the HAZOP approach are listed in Table 3.5 (Rausand, 2011).

Table 3.5: HAZOP framework (Rausand, 2011)

Step	**Planning/preparation phase**
1	Plan and prepare: • Define objectives and limitations • Establish a deliberately balanced HAZOP team • Describe the system (divide into sections, choose major elements for analysis) • Provide background information and data (layout drawings, operation procedures, etc.)
	Execution phase
2	Identify possible deviations: • HAZOP team agrees on the purpose and normal state of the system section • Use guidewords to guide the team into identifying process deviations

3	Identify causes of deviations: • Identify possible causes of deviations
4	Determine consequences of deviation: • Identify possible consequences of deviation
5	Identify existing barriers (safeguards): • Identify safeguards related to the deviation (HAZOP team must be familiar with the existing safety barriers already incorporated in the system)
6	Assess risk: • Estimate the probability and severity, calculate the risk priority number (RPN) • Risk related to each deviation evaluated
	Post work/follow-up
7	Propose improvements: • Propose improvements • Appoint responsible person • Possible comments
8	Report the analysis: • Prepare the HAZOP report from the analysis

3.3.2 Failure Modes and Effects Analysis (FMEA)

Failure Modes and Effects Analysis (FMEA) was one of the first systematic techniques for failure analysis of technical systems. It is a relatively simple technique, which involves assessing the effect of each component part failing in every possible mode in a system. The analysis is carried out for each component in the system to identify and describe all the failure modes, failure causes, and failure effects (Smith, 2011). FMEA is mainly used in the design phase for identifying and analyzing potential failures, but also when identifying parts of the system that should be improved in order to meet today's requirements regarding maintenance, reliability or safety (Rausand, 2011).

The analysis can be carried out by a single person or by a whole team, depending on the complexity of the system. The method does not require any deep analytical skills, but it requires an understanding of the system, its application, and operational and environmental conditions (Rausand, 2011). Therefore, the functions and their performance requirements for each component in the system should be understood and discussed by the FMEA team. Because FMEA originally was made for reliability engineering, the analysis will also cover failure modes that have little or no relevance for the system risk when using it for risk analyses (Smith, 2011).

The FMEA approach is adapted from "Risk Assessment" by Rausand and divided into seven steps (Rausand, 2011), and modified by applying the three main phases (Stangeland, Lundon and Skogvang, 2018). The steps in the FMEA approach are listed in Table 3.6 (Rausand, 2011).

Table 3.6: FMEA framework (Rausand, 2011)

Step	Planning/preparation phase
1	Plan and prepare: • Organization and planning • Identify objectives and limitations • Choose a FMEA team • System description • Provisions of background information
	Execution phase
2	Carry out system breakdown and functional analyses • Define main functions of the system & specify the function performance criteria • Describe operational modes of the system • Break down the system into subsystems that can be handled effectively (for example by establishing a hierarchical structure)
3	Identify failure modes and causes • Identify failure modes • Determine causes of failure • Describe how to detect failure
4	Determine the consequences of the failure modes • Identify consequences of the failure mode on local system levels
5	Assess the risk: • Determine and classify the frequency and severity of the failure mode, and calculate the RPN
	Post work/follow-up
6	Suggest improvements • Possible actions to correct the failure and restore the function or prevent serious consequences are then recorded
7	Report the analysis: • Prepare the report from the analysis • Summarize both the process and the results in an FMEA report

3.4 Status and analysis of gaps

The gaps that have been uncovered for STPA in a workshop-context are the practical implementation of the method and a step-by-step framework (Leveson and Thomas, 2018), as well as deciding whether STPA uncovers more dangerous scenarios than traditional hazard identification methods (Leveson, 2013). Another gap is the implementation of STPA-SEC into the "ordinary" STPA approach in order to identify all potential dangerous scenarios (Leveson and Young, 2013). In Chapter 3, existing step-by-step frameworks and experiences from HAZOP and FMEA, have been used to close some of these gaps (Rausand, 2011). In addition, experts' experiences from previously conducted workshops on hazard identification have been useful for closing the gaps (Stangeland, Lundon and Skogvang, 2018), as well as the theoretical STPA approach provided by Nancy Leveson (Leveson, 2013). Furthermore, the background and the methodology of the STPA-SEC have been studied to be able to integrate the STPA-SEC and the STPA (Yong and Leveson, 2013). Based on the literature found, remaining gaps are considered to be a suggested STPA framework to be used in the workshop, as well as the actual execution of the workshop. In addition, the implementation of STPA-SEC into the "ordinary" STPA, and deciding whether STPA is advantageous are

remaining gaps, as well as creating a final framework for practical implementation and facilitation of STPA. All of these gaps will be treated in the following chapters.

Chapter 4

4. Suggested STPA Framework for the Workshop, and STPA on the "Secure" Function

In this chapter, a framework for practical implementation and facilitation of STPA was prepared and used in the workshop. All of the steps in the workshop are described in detail, and divided into a planning or preparation phase, an execution phase, and a post work phase.

4.1 Suggested STPA Framework

The STPA framework was made before the workshop, and it was mainly based on literature research on STPA, HAZOP and FMEA (Rausand, 2011). The STPA approach described in both "An STPA Primer" (Leveson, 2013) and "STPA Handbook" (Leveson and Thomas, 2018) have been key sources, as well as the frameworks used for HAZOP and FMEA found in "Risk Assessment" (Rausand, 2011). In addition, experts have contributed with presentations that have been held previously on hazard identification. The theoretical foundation on STPA and the STPA approach has been a key resource when planning and executing the STPA workshop (Leveson, 2013). Further on, the study of the HAZOP framework and the FMEA framework was inspiring and gave input on the actual design of the STPA framework that was to be made (Rausand, 2011). In addition, STPA-SEC is included to identify security issues, and thus receive an even better result (Leveson and Young, 2013). Figure 4.1 shows the main elements involved in creating the STPA framework.

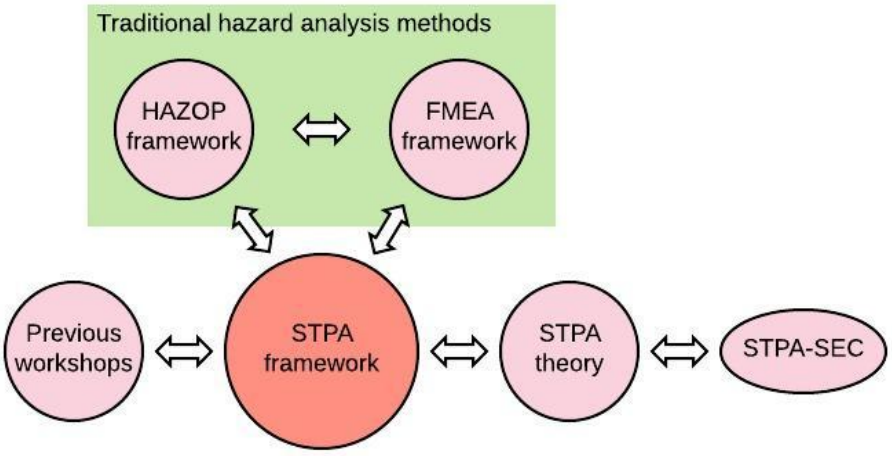

Figure 4.1: Main elements involved in creating the STPA framework

It was decided to divide the framework into three main phases adapted from a previously conducted risk workshop provided by IFE (Stangeland, Lundon and Skogvang, 2018): Planning phase, execution phase, and post work or follow-up. It was also decided to include the STPA-SEC in the suggested framework for STPA, but not to include it in the actual

workshop because of a limited amount of time. As explained in Chapter 3, the STPA-SEC is included in the framework by identifying the system-level security requirements, the unsecure control actions, and the final security requirements (Leveson and Thomas, 2018). Except from that, the STPA is performed in the same way as usual (Leveson and Young, 2013). The STPA framework ended up with ten main steps in total, which are listed in Table 5.1

Table 5.1

Step	Description
Planning/preparation phase	
1	Define purpose of the analysis: • Define objectives • System description • Set system boundary
2	Choose a balanced STPA team: • Establish a deliberately balanced STPA team • Choose a facilitator
3	Describe the system with a control structure: 1. Identify model elements 2. Identify each model element's responsibilities 3. Identify control relationships 4. Identify control actions 5. Develop process model description 6. Identify process model variables 7. Identify process model variable values 8. Identify feedback providing PMV values 9. Check functional control structure model for completeness
4	Provide necessary documentation: • Prepare information on the system, the case study, and the method that will be applied • Prepare a figure that shows the system operation • Prepare a control loop • Prepare a presentation for the workshop • Make a list of necessary guide words
5	Identify system-level accidents, system-level hazards, system-level safety constraints, and system-level security constraints: • Identify system-level accidents • Identify system-level hazards • Identify system-level safety constraints • Identify system-level security constraints • Refine hazards
Execution phase	
6	Identify unsafe/unsecure control actions: • Carry out system breakdown- analyze for each loop • Use guide words for each control action to identify unsafe control actions • Use guide words for each control action to identify unsecure control actions • Structure the answers in an observation form • Refine UCAs by considering safety barriers
7	Identify dangerous scenarios and causal factors: • Identify dangerous scenarios by guide words • Identify causal factor for each scenario • Summarize results- what is critical?
Post work/follow-up	

8	Identify safety/security requirements: • Formulate safety requirements for each UCA • Formulate security requirements for each UCA • Refine safety requirements • Refine security requirements
9	Suggest improvements: • Direct feedback from participants • Survey • Update control loop • Other comments
10	Report the analysis: • Prepare the report from the analysis • Summarize both the process and the results in an STPA report

4.2 Preparation Phase

4.2.1 Step 1: System Conceptualizing

The system conceptualizing is a process where the control loop is formed, and changes are made continuously. The first control loop was made with a lack of knowledge on the system, and without consulting with experts on the field. The control loop had to be updated when missing, technical details to the system were discovered by experts in Bane NOR. Furthermore, another version of the control loop was created after completing the workshop, and new insight into the system's functions was gained. The control loop was then further developed after input from experts, as well as by studying the system and the responsibilities of each actor in detail once again.

The first version of the control loop was made before the workshop, and it is shown in Figure 4.2. It is a simple control structure that shows the relationship between the different components in the "secure" function, as well as five loops that indicate how the communication takes place.

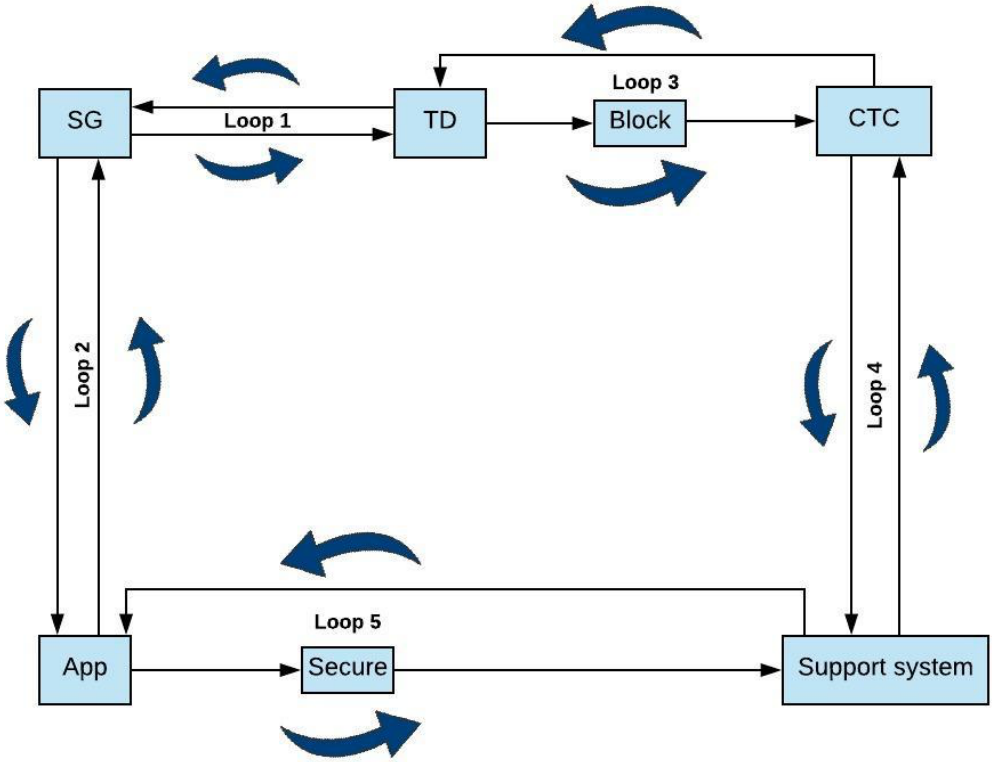

Figure 4.2: First version of the control loop

The second version of the control loop was made after the workshop, and it is shown in Figure 4.3. It was based on input from experts that attended the workshop, and changes made were:

- SuS and CTC switched places to match Figure 2.5, which shows the steps in the "secure" function
- Communication tools were added to the control structure
- Another loop between App and QR code was included
- A loop that goes from start to end in the control loop was added (contains loop 2, loop 3, loop 5, and loop 6)

- Command and feedback lines were added, as well as information exchanged between the components

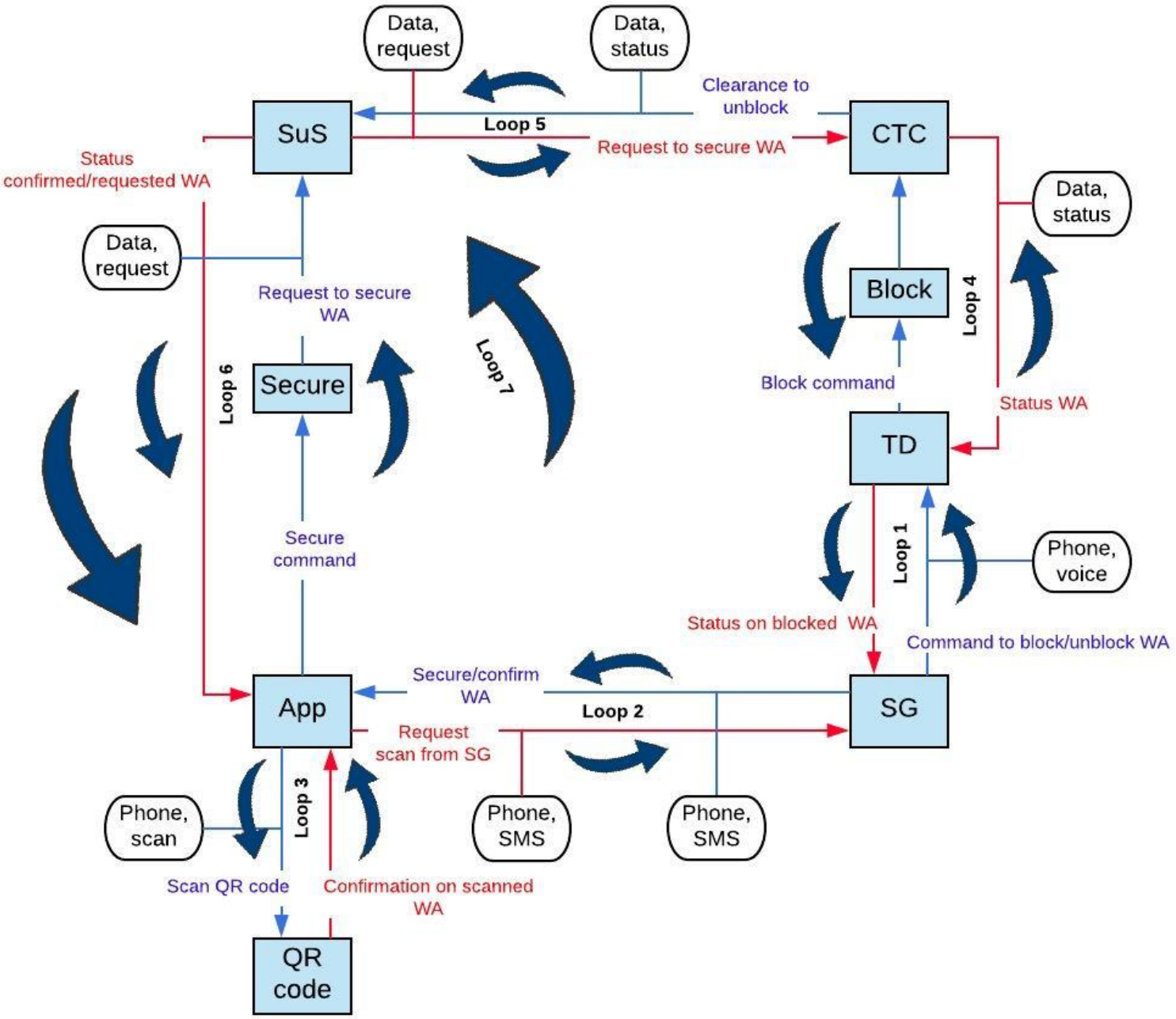

Figure 4.3: Second version of the control loop

4.2.2 Step 2: System-Level Accidents (SLA), System-Level Hazards (SLH), and System-Level Safety Constraints

System level accidents, system level hazards, and system level safety constraints were all defined before the workshop since the accidents were already given as Bane NOR's top events. From the top events, the most relevant accidents for the "secure" function were chosen for further study, which are marked with red in Table 4.2 (Jernbaneverket, 2013).

System-level accidents (SLA)

Table 4.2: System-level accidents

Accident ID	Top event (hazard)	Covers the following single events
SLA-1	Derailment	Failure in rolling stock, superstructure, slippages, over speed, and derailment of dangerous goods
SLA-2	Collision train-train	Collision between trains and other rail transport like work machines etc.
SLA-3	Collision train-object	Collision between trains and various objects on an open stretch and in tunnels: rockslide, animals, road traffic vehicles, tractors or similar objects that randomly have ended up on the line (not at a level crossing).
	Fire	Fire in trains, fire along the tracks, fire in tunnel equipment and explosion, which affect the passengers and train staff.
	Passengers injured on platform	Passengers injured when boarding and disembarking straight and curved platforms, level crossing to platforms in the middle. Also includes events like passengers falling out through doors while the train is moving and passengers injured by trains.
	People injured in level crossing	A train collide into a person or a road traffic vehicle at a level crossing
SLA-4	People injured in and beside the railway tracks (including "electrical safety")	A train collide into a person along the railway track, in contact with high voltage

As seen from Table 4.2, four system-level accidents were classified as relevant to the "secure" function.

System-level hazards (SLH)

The system-level hazards were the hazards that were most likely to happen, and they were found by directly studying the control loop and the SLAs. A total of five system-level hazards were identified, and are listed in Table 4.3.

Table 4.3: System-level hazards and related accidents

Hazard ID	SLH	Related to accident
SLH-1	Miscommunication between SG and TD	SLA-2, SLA-3

SLH-2	Maintenance workers do work on an unsecured work area	SLA-1, SLA-3, SLA-4
SLH-3	TD unblocks the WA prematurely	SLA-1, SLA-2, SLA-3, SLA-4
SLH-4	TD blocks the WA too late	SLA-1, SLA-3
SLH-5	Miscommunication between App and SuS	SLA-2, SLA-3

System-level safety constraints (SLC)

For each SLH, associated SLCs were identified. Table 4.4 shows the associated SLCs to all the SLHs, and a total of nine SLCs were found on a high level.

Table 4.4: System-level accidents and the associated system-level safety constraints

SLH		**SLC**	
SLH-1	Miscommunication between SG and TD	SLC-1	TD must hold the line until correct information is achieved.
		SLC-2	The TD cannot move on to the next step in the procedure before contact between SG and TD is reached.
SLH-2	Maintenance workers do work on an unsecured work area	SLC-3	TD must physically confirm that the WA is secured before maintenance workers start performing work there.
		SLC-4	CTC must confirm that the WA is secured and ready for maintenance work.
SLH-3	TD unblocks the WA prematurely	SLC-5	The TD must wait for the order from SG to unblock the WA
SLH-4	TD blocks the WA too late	SLC-6	The TD must block the WA right after the order is given from the SG
		SLC-7	An alarm must go off if the TD has not blocked the WA after a given amount of time from the order is given
SLH-5	Miscommunication between App and SuS	SLC-8	The CTC must check if the information given to the SG through the SuS and App is correct
		SLC-9	The SG must be alarmed if the information given from the App and SuS does not match the information given from the CTC.

4.2.3 Step 3: Controller Responsibilities and Process Models (PM)

The controller responsibilities and process models were also defined before the workshop because the component responsibilities had already been defined by Bane NOR, while the input to the different components in the control structure were defined by studying the control structure and the responsibilities of each controller. The table of the responsibilities and process model of each controller can be found in Appendix C.

Based on this new information, a third and final version of the control loop was made, which included responsibilities and process models for each controller. Figure 4.4 shows the third version of the control loop, which was also the control structure that was used in the workshop together with Figure 2.5 for the analysis process.

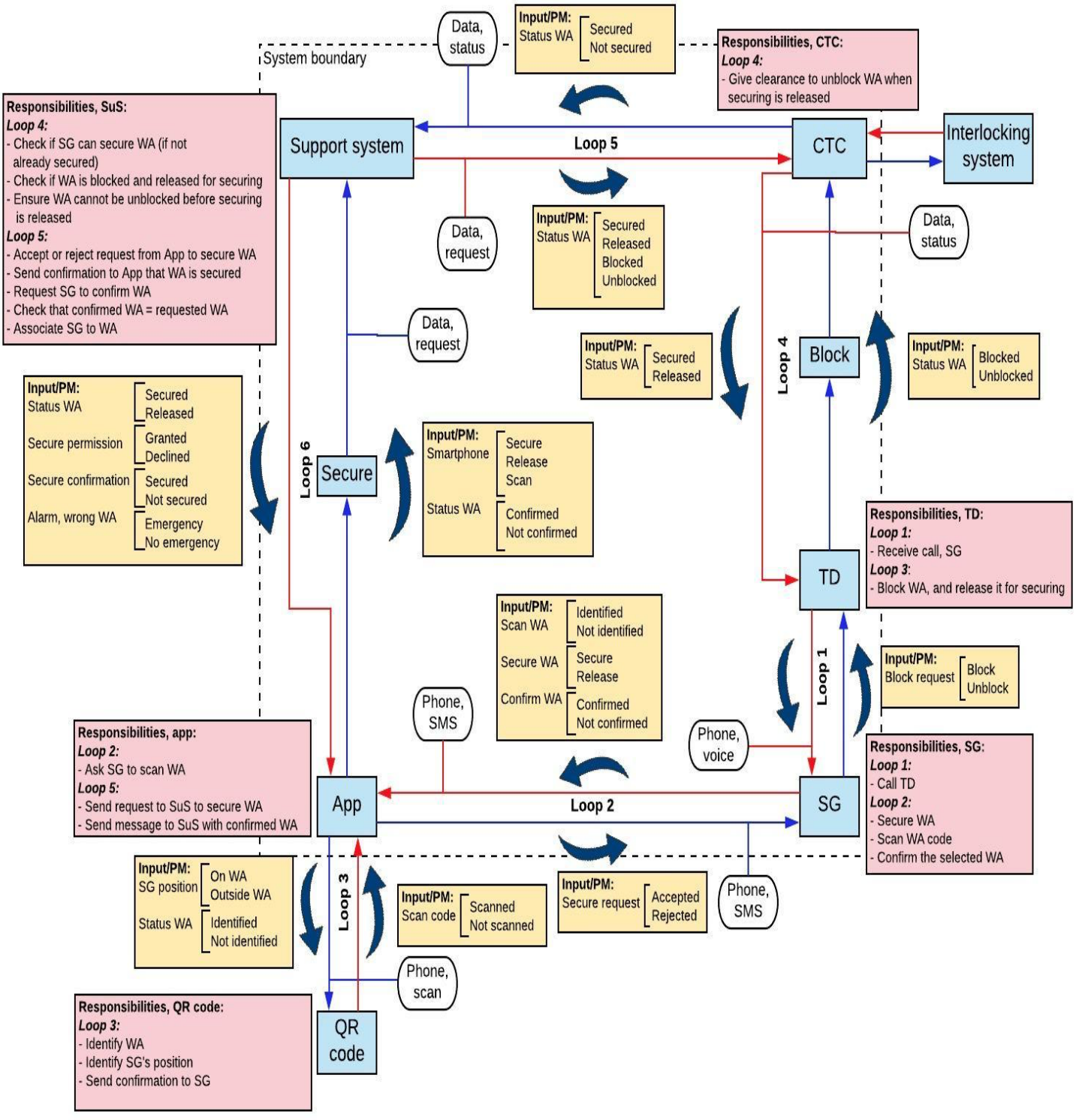

Figure 4.4: Third version of the control loop

4.3 Execution Phase

For the execution phase, an STPA team was chosen to participate in the workshop. The workshop was held in Bane NOR's offices, and it consisted of experts on hazard identification, the project leader, and other engineers that had participated on the previously arranged workshops on the same system.

4.3.1 Step 4 and 5: Unsafe Control Actions (UCA), Dangerous Scenarios and Causal Factors

The unsafe control actions (UCA) and the causal factors were the main focus in the workshop since the UCAs are equivalent to the hazards in the other traditional hazard identification methods, as well as causal factors are identified there too. The dangerous scenarios were defined by using the guidewords provided by Leveson (Leveson, 2013), and the result was the UCAs. The guide words were used in the workshop to initiate the brainstorming process among the participants, and helped with identifying hazards and UCAs. The guide words used in the workshop were (Leveson, 2013):

1. An action is not provided
2. An action is provided, but not followed
3. An unsafe action is provided
4. An action is provided too early
5. An action is provided too late
6. An action is provided in wrong sequence
7. An action is provided too long
8. An action is provided too short

It was decided before the workshop that there would be both a facilitator and an observer. The observer was set to be the student, while the facilitator was set to be the supervisor. The reason for this was that in order for the student to be able to have all questions answered, all focus had to be on observing and taking notes during the workshop. Consequently, the supervisor was chosen to be the facilitator since the supervisor was already familiar with the system, as well has having the knowledge and experience needed.

In order to be able to make notes during the workshop, an observation form had to be prepared beforehand, and the complete observation form can be found in Appendix C. The purpose of the form was to cover the most important aspects of the workshop, as well as identifying which information each of the experts in the workshop contributed with. The QR code was not included in the observation form since it was added to the control structure later

in the process, and therefore the analysis of the QR code was a part of the post work. Table 4.5 shows an example of how the observation form was used through the SG.

Table 4.5 Observation form

Role	Corporation/ Expert	Responsibility (control action)	Scenarios	Input	UCA	Mitigation/ safety barrier	Remaining UCA
SG	Bane NOR (technical experts): BN Safetec (experts on hazard identification and risk analysis): ST IFE (experts on STPA and hazard identification): IFE	Call TD	SG does not call TD SG calls, TD does not respond Call results in unsafe situation SG calls too early SG calls too late SG hangs up too soon	Block status (from TD)	SG hangs up too early, and does not get confirmation of block status (BN) SG makes the call before arriving at the WA (IFE)	Support system checks status of blocking and confirms release (BN)	…

A total of forty-five UCAs were identified in the STPA workshop, and a full table of all the dangerous scenarios, UCAs, and causal factors can also be found in Appendix C.

4.4 Post Work

4.4.1 Step 6: Remaining UCA

To avoid excessive UCAs, the remaining UCAs were identified after the workshop by considering the requirements specification given by Bane NOR for the entire system (Sivertsen, 2014). In Table 4.6, the first remaining UCAs identified are listed when considering safety barriers. The remaining UCAs were eight in total, and the rest of the table can be found in Appendix C.

Table 4.6: UCAs, safety barriers, and remaining UCAs

ID	UCA	Already existing safety barrier	Planned safety barrier, SC (from Bane NOR/experts)	Remaining UCA

1	SG does not call TD	TD is in a "deadlock", and no securing or blocking is possible until the call is made		
2	SG makes the call before arriving at the WA			1. SG makes the call before arriving at the WA
3	SG loses data connection when calling TD		It is not possible to continue to the next step in the "secure" process until the connection is re-established	

4.4.2 Step 7: Safety Constraints

The safety constraints associated to each of the remaining UCAs were identified after the workshop had been completed, and therefore they were also a part of the post work. A total of nine safety constraints were identified for the remaining UCAs, and they are listed in Table 4.7.

Table 4.7: Remaining UCAs and the associated safety constraints

UCA ID	**Remaining UCA**	**SC ID**	**Safety constraint (SC)**
1	SG makes the call before arriving at the WA	1	Information about which WA the SG shall secure is provided in advance
		2	The SG must state which WA to secure, and the SG's position must be checked to see if the information stated is correct (for example, a positioning system can be used)
2	SG pushes the wrong button on the App when securing	3	The App requests a confirmation of the function selection from the SG
3	SG scans the QR code two times in a row, such that the confirmation is given too early	4	SG cannot confirm the WA before SuS has requested a confirmation
		5	There must be a minimum time between scanning the WA for securing and confirming the WA
4	Wrong information on the status of the WA is given to the SG	6	All input data must be appropriately validated through data integrity requirements
5	The app distorts the information from SuS, even though it is transferred correctly	6	All input data must be appropriately validated through data integrity requirements

6	WA requested has been covered earlier, and no confirmation is given	7	Temporary status functions for the WA must be used until the final confirmation is given
7	A confirmation that the WA can be secured/released is sent too late from the SuS to the App	8	The App and the SuS must monitor each other and thus function as mutual technical barriers
8	SuS associates SG to WA, but SG switches WA without registering in the App	9	The SG must always register in the App when switching WAs, and if it is forgotten the SuS must notify the SG through the App (for example, a positioning system can be used)

4.4.3 Step 8: Evaluation and Comments to the Workshop

After the workshop was completed, the participants of the workshop gave direct feedback to both the facilitator that was in charge of the execution of the workshop, and the observer that had done the preparations and made the presentation. Suggestions to technical changes to the control loop were also made, but they are mentioned under system conceptualization. In addition, a survey was sent out afterwards to all the participants in order to receive as much feedback as possible. The survey gave the participants the possibility to rate different things related to the workshop on a scale from one to five, and all the questions and answers can be found in Appendix D.

Overall, the feedback from the different experts was positive. Experts from Bane NOR described the control loop as well-structured, and they thought that it is was suited for people familiar to the system only. The analysis process in the workshop was described as similar to a HAZOP, but a HAZOP is better suited for analyzing the physical parts of the system. Experts from IFE agreed on the good quality of the control loop, and they also stated that knowledge to the system is crucial when performing an STPA. The preparations were said to be good, and that the facilitator did a good job by opening up for discussions throughout the workshop. Furthermore, experts from Safetec agreed on what was said about the control loop and preparations made beforehand. In addition, it was said that it might be a disadvantage with a high level of knowledge, but that it also opens up for good discussions. The complete table of all feedback given can be found in Appendix C.

Summary

The UCAs and causal factors can be used to see how well the STPA does compared to the traditional hazard identification methods. To summarize the most important findings;

- A total of 44 UCAs were identified without considering already existing or planned safety barriers

- A total of 8 UCAs were identified when considering already existing or planned safety barriers
- A total of 9 safety constraints relating to the 8 remaining UCAs were identified

In Chapter 4, more of the gaps mentioned in Chapter 3 have been closed. A suggested framework for practical implementation and facilitation of STPA has been created, and the STPA-SEC approach has been implemented into the "ordinary" STPA approach. In addition, the workshop has been conducted and evaluated, such that a final framework can be made. A remaining gap is now to confirm or reject Leveson's claim about STPA being advantageous to use at complex systems, and that it identifies more hazards than the traditional hazard identification methods. Another gap that must be treated is the final framework for practical implementation and facilitation of STPA. It is not possible to confirm or reject Leveson's claim until the results from the STPA workshop has been compared to the results of some traditional hazard identification methods, which is done in the next chapter.

Chapter 5

5. Comparison of Risk Assessment Methods: STPA vs. HAZOP and FMEA

Previously in the SafeT project, a HAZOP workshop and an FMEA workshop have been conducted in order to identify possible dangers in the system for securing a work area, and for the "secure" function in the same system. In this Chapter 5, the results from the STPA workshop is compared to the previous findings to see if STPA is advantageous to use at complex systems, and if it reveals more hazardous scenarios than the HAZOP and FMEA. This chapter therefore handles the gap concerning Leveson's claim about STPA being advantageous to use over traditional hazard identification methods.

5.1 Background

In many cases, STPA has proven to theoretically provide a wider scope and to identify more hazards compared to the traditional hazard identification methods (Leveson, 2011). Nevertheless, STPA is a relatively new hazard identification method to the railway industry in Norway, and it is therefore of interest to study similarities and the main differences between STPA and other traditional hazard identification methods. During the comparison, in order to have something that is measurable, the main focus has been on the number of hazards identified in the "secure" function in each of the methods.

In the SafeT project, some workshops have already been conducted where the hazard identification methods used have been HAZOP, SysML, and FMEA (Gran, Karpati and Hauge, 2018). In this thesis, HAZOP and FMEA have been selected for the comparison between STPA and traditional hazard identification methods. The reason for this is that the reports from the HAZOP workshop and the FMEA workshop have already been finalized, while the report on the SysML workshop is still being prepared. In addition, SysML is a more advanced method, which might have complicated the comparison. Similarities and differences regarding the approaches and the execution of the workshops are discussed as well.

5.2 STPA vs. HAZOP

In the HAZOP, the hazards are equivalent to the UCAs identified in the STPA, while the measures are equivalent to the safety constraints identified. In the HAZOP, hazards related to RAM were also identified, but in this thesis only the hazards related to safety are of interest. Therefore, only the hazards related to safety were taken into consideration in the comparison between the STPA and the HAZOP. In the HAZOP report, only main functions 1, 2, 5, and 6

are studied because the other functions are not considered as safety critical (Opsahl, Solibakke and Skogvang, 2018).

In the HAZOP report, the hazards are listed as both general hazards and hazards found in the "secure" function (Opsahl, Solibakke and Skogvang, 2018). For this comparison, only the hazards that were identified for the "secure" function were of interest, and the results are listed in Table 5.1. A pre-condition for performing the main function "secure" in the HAZOP report was that the train dispatcher had blocked the work area already (Opsahl, Solibakke and Skogvang, 2018).

5.2.1 Results from the HAZOP Report

Table 5.1: Hazards, causes and measure identified in the HAZOP for the "secure" function (Opsahl, Solibakke and Skogvang, 2018)

Main function: Secure				
Hazard ID	**Hazard description**	**Cause**	**Measure ID**	**Measure description**
H-0025	Main safety guard prevented from using the GSM-R unit	Loss of power, spurious log out, etc.	M-0025	The OSS power of authority must be clearly defined. OSS actions may be safety critical
H-0026	Corruption of the QR code	Someone tries to harm the railway infrastructure and uses a malicious QR code to get access to the system	M-0026	Use of dual coding. Use of GPS may also be a "supporting tool" for the identification of the location of the SG (the GPS is accurate enough to give the precise location)
H-0027	Main safety guard scans the wrong code	On stations for example, there may be a number of tracks and working areas. The safety guard may therefore scan the wrong code.	M-0027	
H-0028	Voice communication is not functioning as intended, or it is too noisy to talk to the TD	If the voice communication is not functioning, the communication can be performed via SMS	M-0028	There should be a procedure for all communication in abnormal situations

In order to be able to compare the two hazard identification methods, the hazards identified in the HAZOP and the associated UCAs from the STPA are compared. Because safety barriers are not considered in the HAZOP, the hazards are compared to the UCAs identified in the STPA before considering safety barriers. The hazards and associated USAs are listed in Table 5.2.

Table 5.2: Hazards identified in the HAZOP, and the associated UCAs in the STPA

HAZOP		**STPA**	
Hazard ID	**Hazard description**	**UCA ID**	**UCA description**
H-0025	Main safety guard prevented from using the GSM-R unit	3	SG loses data connection when calling TD
		17	Data connection is lost, but TD still has voice connection
H-0026	Corruption of the QR code	28	The QR code has been counterfeited such that it is not possible to identify the WA
		29	It is not possible to scan the QR code because of vandalism
		31	A copy of the QR code has been made such that the code can be scanned without the SG physically being at the WA
H-0027	Main safety guard scans the wrong code	13	SG scans the wrong QR code
H-0028	Voice communication is not functioning as intended, or it is too noisy to talk to the TD	15	TD is not available when SG calls
		16	TD does not respond, and SG secures WA anyways
		17	Data connection is lost, but TD still has voice connection
			TD cannot hear SG because of noise

Table 5.2 shows that the HAZOP identified four hazards, which are equivalent to nine of the UCAs identified in the STPA. In addition, the HAZOP identified a new UCA concerning errors in the voice communication because of too much noise on the work area.

5.2.2 Similarities and Differences between STPA and HAZOP

The hazards identified in the HAZOP are similar to the UCAs identified in the STPA, but the UCAs identified in the STPA are more detailed than the hazards in the HAZOP report. In addition, the hazard descriptions in the HAZOP report were vague and did not contain detailed information (Opsahl, Solibakke and Skogvang, 2018). Nevertheless, the measures given in the HAZOP are more detailed than the safety constraints formulated in the STPA (Opsahl, Solibakke and Skogvang, 2018).

5.3 STPA vs. FMEA

In the FMEA, the failure modes are equivalent to the UCAs identified in the STPA, and the barriers are equivalent to the safety constraints. The possible failure modes in the FMEA report are loss, partial loss, delay, and corruption, while the possible actors are sender, message medium, and receiver (Gran, 2018).

The "secure" function is the subject for the analysis in the FMEA report, and a night time session is considered, which normally lasts four hours (Gran, 2018). Only the relevant information from the FMEA report is included, which is shown in the table below as failure mode, cause, and barrier. It is assumed that if the App does not work, the support system can be used instead with simple SMS messages. Furthermore, the data protocols are not defined for communication between different agents and actors. In addition, it is assumed that all alarms are sent to the right users (Gran, 2018). Table 5.3 shows the results from the FMEA report (Gran, 2018).

5.3.1 Results from the FMEA Report

Table 5.3: Failure modes, failure causes, and barriers identified in the FMEA for the "secure" function (Gran, 2018)

Main function: Secure				
Failure ID	**Function**	**Failure mode**	**Failure cause**	**Barrier/mitigation**
1	Initiate call	Unable to call/interruption, corruption	No signal	Ensure good signalling condition with GSM-R at all relevant WAs. Assumption is GSM-R is available
			Wrong number-faulty contact information	Assume that correct contact information for TD (the role and not person) of relevant WA is available on phone or informed to SG
2	Answer call	Unable to answer call/interruption, corruption	Occupied/unavailable due to high work load or sudden events	GSM-R functionality
3	Select function call	Unable to select (loss)	App freeze	Requirements on phone and app robustness to accommodate all potential environmental condition expected to experience Possibility to use secondary device as fall-back

4	Select function call	Wrong function selected	Unintended selection of wrong function	Assume display in app of user selection Confirm selection on all critical functions
5	Check authority access for SG at WA	Erroneously denied access	Access authority not updated	Unsure how access is managed and granted
			Access authority not updated- certificates out of date, training not given	Unsure how access is managed and granted
6	Inform SG that next stop is to scan code	No display in app of expected user action	Software/hardware failure	
7	Scan and interpret physical sign	Unable to scan	Sign damaged/polluted	
8	Sign "send" information of WA	Damaged sign	Legal but erroneous information (dirty sign)	Dual coding where sign both has text and code and SG can confirm that correct WA is secured. Not specified in the model how the dual coding is used by SG and App to confirm correct location
		Sabotaged message/sign	Legal but other information carrier (scan picture of sign or another phone)	Procedures- SG break procedures

Again, in order to be able to compare the two hazard identification methods, the associated UCAs are put together with the failure modes identified in the FMEA report in Table 5.4. Also in this case, the failure modes in the FMEA are compared with the UCAs identified in the STPA before considering safety barriers.

Table 5.4: Failure modes identified in the FMEA and the associated UCAs from the STPA

FMEA		**STPA**	
Failure ID	**Failure mode**	**UCA ID**	**UCA description**
1	Unable to call/interruption, corruption (initiate call)	1	SG does not call TD
		3	SG loses data connection when calling TD
2	Unable to answer call/interruption, corruption	15	TD is not available when SG calls
		16	TD does not respond, and SG secures WA anyways
3	Unable to select call function		
4	Wrong function selected	9	SG pushes the wrong button on the App when securing

5	Denied authority access for SG at WA	44	SuS associates SG to WA, but SG switches WA without registering on the App
		45	SuS associates SG to wrong WA
6	No display in app when SG is informed that next stop is to scan code	22	The App does not ask the SG to scan the relevant WA
		23	The request from the App to the SG is not sent
7	Unable to scan and interpret physical sign	12	SG does not scan the QR code
		13	SG scans the wrong QR code
		14	SG scans the QR code two times in a row, such that the confirmation is given too early
8	Sign "send" wrong information of WA	28	The QR code has been counterfeited such that it is not possible to identify the WA
		29	It is not possible to scan the QR code because of vandalism

Table 5.4 shows that the FMEA identified eight failure modes, while the STPA identified 14 associated UCAs. In addition, the FMEA identified one new failure mode concerning not being able to select the call function.

5.3.2 Similarities and Differences between STPA and HAZOP

In both of the hazard analyses the focus was the "secure" function only. Many of the same failure modes or UCAs were covered in the STPA and FMEA reports, even though FMEA had one extra failure mode. Nevertheless, less failure modes and UCAs were found in total in the FMEA both when considering all of the forty-five UCAs and the sixteen UCAs after mitigation. Another difference, was the level of detail on the barriers and mitigation in the FMEA compared to the STPA. The barriers in the FMEA included much more details and had a higher complexity.

5.4 Summary

To summarize the findings from the comparison of the STPA and the traditional hazard identification methods, HAZOP and FMEA:

- 4 hazards were identified for the "secure" function in the HAZOP report, which is equivalent to 9 of the UCAs identified in the STPA
- 8 failure modes were identified for the "secure" function in the FMEA report, which is equivalent to 14 of the UCAs identified in the STPA
- Many of the same UCAs were covered in all of the three hazard identification methods
- Two new UCAs were identified in the HAZOP and in the FMEA

Based on the comparison, the STPA identified more hazards than the traditional hazard identification methods, even though there were two new hazards identified in the HAZOP and

in the FMEA. There were forty-five UCAs identified in total in the STPA when not considering the safety barriers, which are far more than identified in the HAZOP and the FMEA. Consequently, STPA is advantageous to use for complex systems, and Leveson's claim is confirmed, which means that another gap has been closed. The last gap is to create the final framework for practical implementation and facilitation of STPA, which is covered in the next chapter.

If the STPA was expanded to include STPA-SEC, it is likely that more UCAs associated with ICT security could have been identified. In that case, controllers or components that affect or are involved in the computer system in the "secure" function would have been in focus, and not the human controllers (Friedberg et al., 2017). Components in the "secure" function that would have been of interest in an SPTA-SEC are the support system, the CTC system, the App, and the QR code. The STPA-SEC focuses on security instead of safety, but the approach is similar to the "ordinary" STPA approach, except from some extra steps as explained in Chapter 3. The extra steps would be to identify system-level security constraints, unsecure control actions, and final security requirements (Leveson and Thomas, 2018).

Chapter 6

6. Final STPA Framework for Practical Implementation and Facilitation of STPA

This chapter presents the final framework for the practical implementation and facilitation of STPA after the workshop has been conducted, and thereby closes the last gap concerning the development of a final STPA framework. The framework is then illustrated by inputs, outputs, and main steps, before each of the main steps are described in detail.

6.1 Assumptions Made for Making of the Framework

The basis for the framework created before the workshop and used in the workshop was literature research, as well as input from experts. The final framework for practical implementation and facilitation of STPA was made after the workshop had been conducted, and after the framework suggested in Chapter 4 had been tested. When developing the final framework for practical implementation and facilitation of STPA, experiences gained from the workshop and the feedback from the evaluation process can be considered as well. Consequently, the basis for the final framework have been the previously suggested framework and own experiences from the workshop, as well as feedback and suggested improvements from experts in the evaluation phase. Figure 6.1 shows the main elements involved when making the final framework for practical implementation and facilitation of STPA.

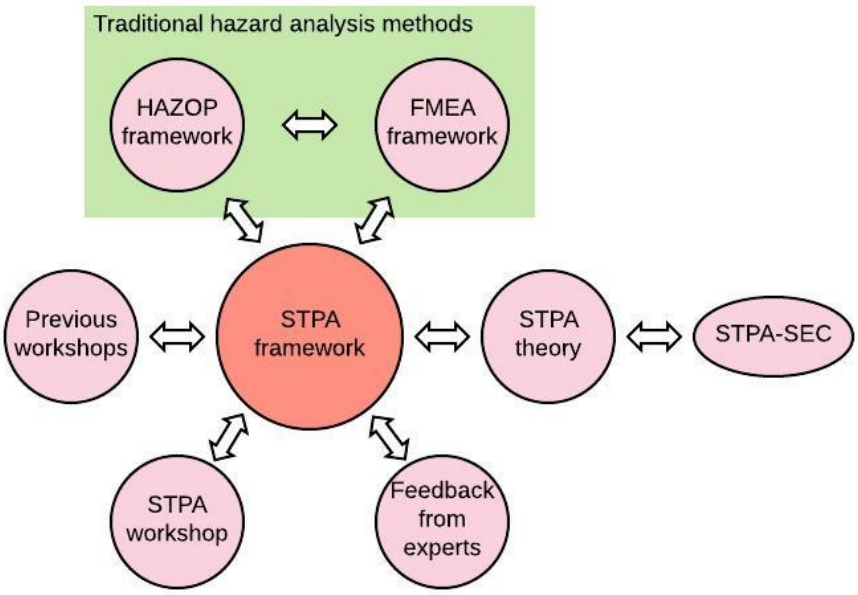

Figure 6.1: Main elements involved in creating the final STPA framework

As seen from the evaluation and feedback given after the workshop (See Appendix C), the experts agreed on many aspects of the workshop and the analysis process. The preparations done beforehand were good and informative, the control loop and the system operation figure used for the analysis were well-structured, and the facilitator did a good job in initiating good discussions. They also agreed on the fact that all the participants must have knowledge about the system in order for the STPA to be successful. On the other hand, it was discussed whether being familiar with the system beforehand could be a disadvantage as well.

The new input and feedback from experts that participated in the STPA workshop does not lead to any major changes in the main steps in the STPA framework, but it contributes to important information that must be mentioned when describing the steps. Therefore, it is decided to keep the STPA steps listed in Table 5.1, and rather make a detailed figure of the final framework, as well as a detailed description of each step.

6.2 Illustration of the STPA Framework

In order to create a well-structured figure of the STPA framework that was easy to follow, it was decided to adapt the design of the frameworks used for the HAZOP and the FMEA (Rausand, 2011), which meant to include inputs, steps, and outputs. Consequently, inputs and output for each step had to be identified, which could be adapted from "STPA handbook" (Leveson and Thomas, 2018) and from the results in Chapter 4. All of the inputs and outputs for each step are listed in Table 6.1.

Table 6.1: Inputs, steps, and outputs in the STPA framework

Input	**Steps**	**Output**
	Define purpose of the analysis	• Defined objectives • Defined goals • System boundary
	Choose a balanced STPA team	• Study team • Facilitator • Project plan
• Information/ documentation on the system	Describe the system with a control structure	• Control loop • Responsibilities/ control actions/ • Process models • Feedback loops
• Data sources • Experience data	Provide necessary documentation	• System description and operation • List of guide words • Presentation
• Experience data	Identify system-level accidents, system-level hazards, system-level safety constraints, and system-level security constraints	• SLA • SLH • SLC (safety/security)

• Experience data • SLH • Control loop • Control actions • Process models • List of guide words • List of safety/security barriers/requirements • Expert judgment	Identify unsafe/unsecure control actions	• List of UCAs (safety/security) for each control action • List of remaining UCAs (safety/security)
• Experience data • Control loop • Control actions • UCA (safety/security) • Process models • List of guide words • Expert judgment	Identify dangerous scenarios and causal factors	• List of dangerous scenarios • List of causal factors for each scenario
• Experience data • UCA (safety/security) • Dangerous scenario • Causal factor • Expert judgment	Identify safety/security constraints	• List of safety requirements for each UCA (safety/security)
• List of UCAs (safety/security) • Control loop • Feedback • Expert judgment	Suggest improvements	• List of relevant improvements • Updated control loop
• List of relevant improvements • Updated control loop	Report the analysis	• STPA report

From Table 6.1, a figure of the STPA framework could be made, which is illustrated in Figure 6.2. Because STPA-SEC is included in the framework, UCA in the figure means both unsafe control actions and unsecure control actions, while SC in the figure are both safety constraints and security constraints.

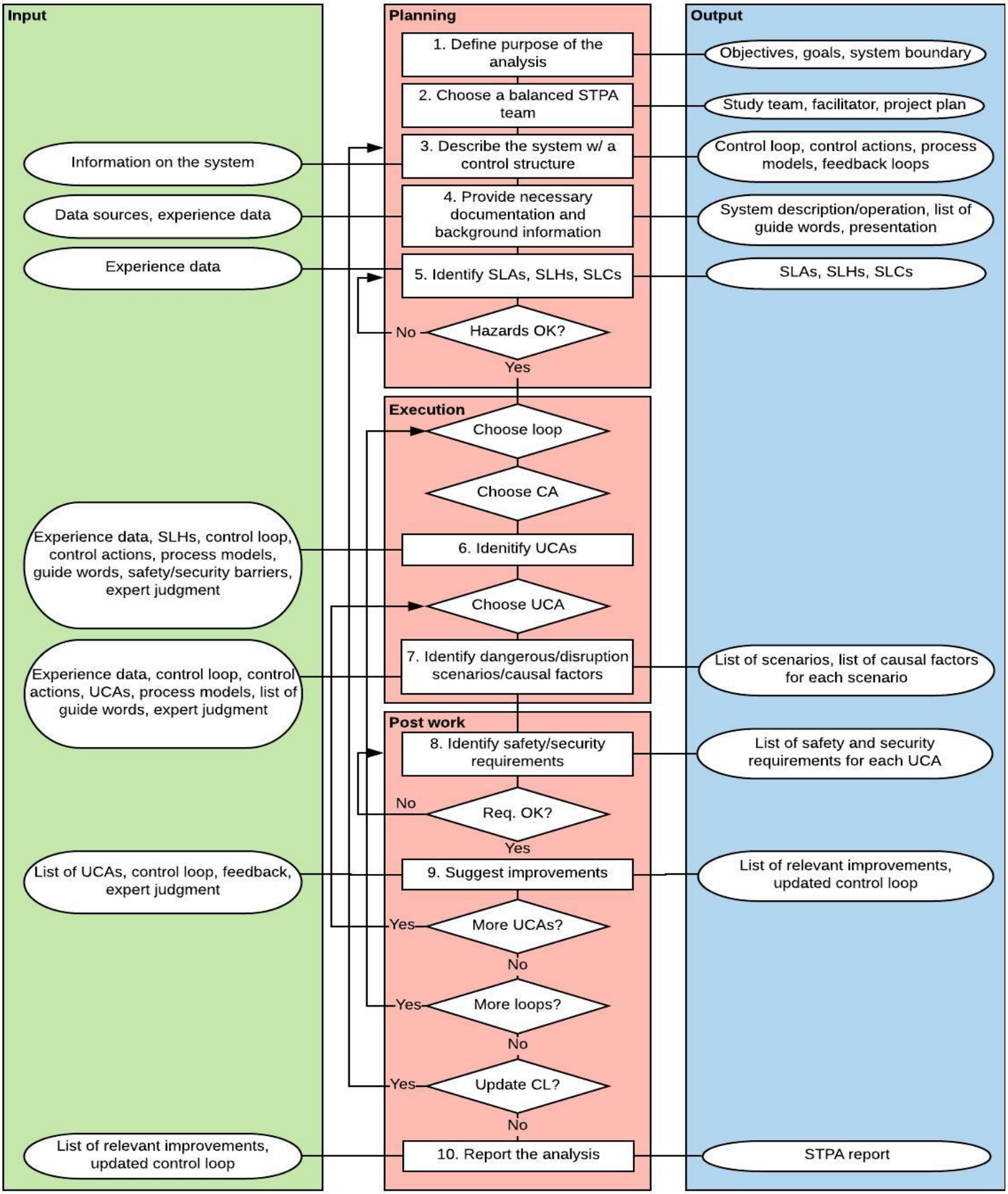

Figure 6.2: STPA framework with inputs and outputs for each step

6.3 Description of Each Step in the STPA Framework

To ensure that the steps are performed correctly, detailed descriptions of all the steps are provided as well. The descriptions of the steps are summaries the steps provided in Chapter 3 and Chapter 4, as well as some input from "STPA Handbook" (Leveson and Thomas, 2018) and "Risk Assessment" (Rausand, 2011). Some aspects from "A systems approach to risk assessment in maritime operations" are also included (Rokseth et al., 2016). In addition, own experiences and feedback given after the workshop have been considered (See Appendix C).

6.3.1 Planning/Preparation Phase

Step 1: Define the purpose of the analysis

The first step in the STPA is to identify the scope and the objectives of the analysis, which is done by considering what the system shall do (purpose), and how the system will do it (method) in order to achieve something (goal) (Young and Porada, 2017).

Step 2: Choose a balanced STPA team

An STPA team should consist of 5-8 team members, and it is important to include different fields of expertise when selecting the team. The selection of team members will depend on the complexity and purpose of the system (Rausand, 2011), and the experts should be involved in an early stage and in the entire process. Ideally, the team members should not have participated in a workshop with the same system earlier to be able to identify as many new UCAs as possible. STPA does not require any deep analytical skills, but it is recommended that all team members have an understanding of the system under consideration, its application, and operational and environmental conditions (Rausand, 2011). In order to have an effective STPA, it is also recommended to include experts that have knowledge about the design and operational part of the system. Experts that have technical knowledge about the system can make sure that excessive UCAs are not included in the results. There must be an STPA leader, who will also be the facilitator in most cases, and who is responsible of leading the workshop and using guide words to trigger valuable discussions (Rausand, 2011). In addition, an STPA secretary is necessary to make notes during the process (Rausand, 2011) and to summarize the results of the workshop by using an observation form (see Appendix C). The STPA team must together agree on the objectives of the workshop before performing the analysis, plan and estimate the time of the study, and decide the style of recording. Figure 6.3 shows an example of the composition of an STPA team.

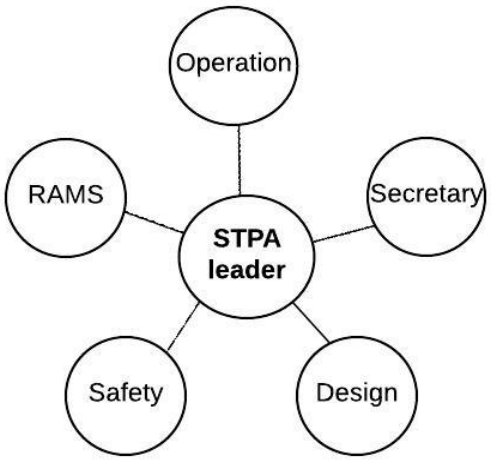

Figure 6.3: Example of an STPA team

Figure 6.3 shows the STPA leader in the middle because he is in control of the communication during the workshop, and the secretary to his right because there is a close collaboration between them (Rausand, 2011). The STPA leader and the STPA secretary shall walk through the presentation together before the workshop takes place. Examples of other experts that may participate in and STPA workshop are operations manager, design engineer, safety engineer, and Reliability, Availability, Maintainability and Safety (RAMS) engineer (Rausand, 2011).

Step 3: Describe the system with a control structure

The control loop should illustrate which components that are included in the system, as well as the relationships between them. The way the control loop is made sets the system boundary and decides the scope of the analysis (Rokseth et al, 2017). The control loop should be made early in the process, before the actual analysis takes place, such that there is plenty of time to review and update it during the process. Whenever using a control loop only for the analysis, it is recommended that all team members are familiar with the system as the control loop alone does not provide enough information to explain the system (see Appendix C). The procedure of making a control loop is described as following (Young and Porada, 2017):

1. Identify all components
2. Identify each component's responsibilities
3. Identify control relationships
4. Identify control actions (CA)
5. Develop process model description
6. Identify process model variables (PMV)
7. Identify process model variable values
8. Identify feedback providing PMV values
9. Check functional control structure model for completeness

The control structure should consist of at least five types of elements, which are controllers, control actions, feedback, inputs and outputs from components, and controlled processes (Leveson and Thomas, 2018). Figure 6.5 shows a generic control loop, which is adapted from "STPA Handbook" (Leveson and Thomas, 2018).

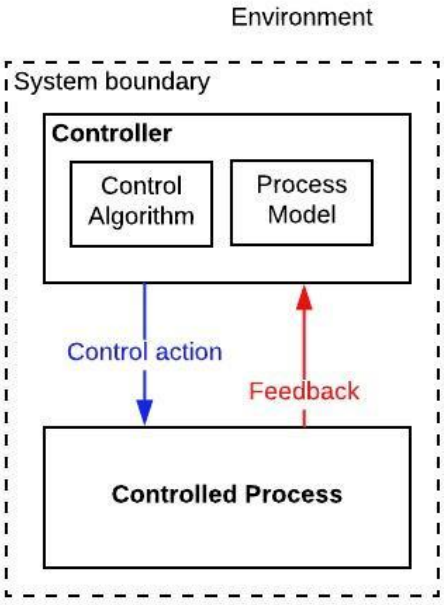

Figure 6.4: Generic model of a control loop (Leveson and Thomas, 2018)

Each controller consists of a control algorithm and a process model. The control algorithm can be defined as the controller's decision-making process, and it determines which control action to provide. The controller performs a control action on a controlled process, which respond by feedback (Leveson and Thomas, 2018). The control action and feedback will be different for the safety focus and the security focus, and therefore additional arrows must be made when implementing STPA-SEC into the control loop. The more information about the system that can be obtained from the control loop, the better the results will be (see Appendix C).

Step 4: Provide necessary documentation and background information

Before arranging the workshop, necessary documentation and background information must be provided to all the participants (Rausand, 2011). These documents should contain information about the system and the system operation, as well as the case study and the method that will be applied in the workshop. These documents must be sent out beforehand such that the participants have plenty of time to prepare, and consequently have an effective meeting. All the important terms and key concepts applied in the analysis must be explained either in these documents or at the start of the workshop. In addition, a list of guide words that will be used in the analysis must listed. Before the meeting, practical information like agenda,

topics introduced at the meeting, number of participants, and how the analysis will be performed should be provided. For hazard analyses it is recommended not to exceed three hours as the participants will get impatient and unfocused (Rausand, 2011). Therefore, it is recommended to rather perform the STPA over multiple meetings if there are large systems to analyze.

6.3.2 Execution Phase

Step 5: Identify system-level accidents (SLA), system-level hazards (SLH), and system-level safety and security constraints (SLC)

The SLA, SLH, and SLC are also identified before the STPA workshop takes place. Structured tables should be used, as well as data sources and experience data. Examples of experience data used in this step are accident data, common hazards, and requirements or safety and security barriers, depending on the system under consideration. Table 6.2 and 6.3 show how the SLAs, SLHs, and SLCs can be structured. There must be separate tables for the safety constraints and the security constraints because the accidents and hazards have different focus. For the safety constraints, the accidents and hazards will be related to safety issues, and for the security constraints, the accidents and hazards will be related to security issues.

Table 6.2: System-level accidents, system-level hazards, and system-level safety constraints

System-level accident	**System-level hazard**	**System-level safety constraint**
SLA-1	SLH-1	SLC-1
		SLC-2
SLA-2	SLH-2	SLC-3

Table 6.3: System-level accidents, system-level hazards, and system-level security constraints

System-level accident	**System-level hazard**	**System-level security constraint**
SLA-1	SLH-1	SLC-1
		SLC-2
SLA-2	SLH-2	SLC-3

Usually, there is one corresponding hazard for each accident identified, while there can be several safety constraints or security constraints for each hazard.

Step 6: Identify unsafe and unsecure control actions

Both the unsafe and unsecure control actions should be identified by using guide words, which is done by breaking down the system into smaller parts or loops, and then analyzing each part or loop. The same generic guide words can be applied to the unsafe and unsecure

control actions. A control action for each component in the system is put together with the guide words to identify unsafe or unsecure control, and thereby unsafe and unsecure control actions. The eight guide words that can be used, which are generic modes of unsafe or unsecure control, are listed below:

1. An action is not provided
2. An action is provided, but not followed
3. An unsafe action is provided
4. An action is provided too early
5. An action is provided too late
6. An action is provided in wrong sequence
7. An action is provided too long
8. An action is provided too short

All components in a loop should be analyzed before moving on to the next loop in the control structure.

Step 7: Identify dangerous scenarios and causal factors

In this step, each part or loop of the control loop is investigated in order to find how the UCAs could possibly occur. Because process models are often involved in dangerous scenarios, they are studied for each UCA.

Structured tables are used to summarize the results, and Table 6.4 shows an example of how an observation form may look like to cover all of the important aspects of an STPA. Two observations forms must be filled out when the STPA-SEC is included in the STPA because the control actions, dangerous scenarios, inputs, UCAs, and barriers will vary depending on whether they are safety-related or security-related.

Table 6.4: Example of observation form for the STPA workshop

Role	**Expert**	**Control action**	**Dangerous scenario (guide words)**	**Input/ process model**	**UCA**	**Safety/ security barrier**	**Remaining UCA**

In this observation form the most important topics are covered, and the results needed for further study are summarized. The different columns are described below:

- Role: The role or the component currently studied in the system.
- Expert: The company or person who comments during the analysis
- Control action: The command or the responsibility performed by a component or controller on a controlled process

- Dangerous scenarios: The scenarios identified when applying guide words to the control actions
- Input/process model: The different states each of the components or controllers can be in
- UCA: Potentially unsafe or unsecure incidents where inadequate control occurs
- Safety/security barrier: The safety or security barriers are measures that prevent the UCAs from happening
- Remaining UCA: The remaining UCAs are the potentially unsafe or unsecure incidents where inadequate control occurs after considering the safety or security barriers

6.3.3 Post Work

Step 8: Identify safety and security requirements

Safety and security requirements are made during the post work or follow-up process, and it is done by considering each causal factor. The UCAs identified are used together with information on how and why they might occur in order to formulate the safety and security requirements (Leveson, 2011). There might be several safety or security requirements for each UCA.

Step 9: Suggest improvements

Possible improvements on the results or the control loop should be discussed and noted by the STPA secretary at the end of the meeting. Suggestions can be made through discussions and through direct feedback from the participants. If there are disagreements regarding the results of the analysis, parts of the system may be restudied by going back to step 6. It must be followed up that the suggested improvements actually are implemented.

Step 10: Report the analysis

When making the STPA report, it can be useful to include the observation form, which includes all the topics discussed during the meeting. In addition, the report should contain the unsafe and unsecure control actions that were discovered, as well as the remaining unsafe and unsecure control actions that were identified after considering the already existing safety or security barriers, as well as the safety and security requirements. The report should be sent out to all the participants for them to review and make comments if there are any disagreements regarding the results or things that were said. Especially important is the review process regarding the safety and security requirements.

Chapter 7

7. Conclusions, Discussion, and Recommendations for Further Work

7.1 Summary and Conclusions

In this thesis, an STPA workshop was planned and executed based on the theoretical foundation of STPA and other risk assessment methods, as well as experience data from both a HAZOP workshop and an FMEA workshop. The system that was analyzed in the STPA workshop was the "secure" function in the system for securing a work area when maintenance is performed there. The STPA team consisted of experts on hazard identification, the project leader, and engineers that had participated on previously arranged workshops on the same system. The results of the workshop were the unsafe control actions identified, which could be compared with the dangers identified in the HAZOP workshop and the FMEA workshop conducted earlier in the SafeT project. The STPA-SEC methodology was studied in order to figure out how to include dangers associated with ICT security in an STPA review, and then implemented into a suggested STPA framework that was followed throughout the workshop.

A total of 44 unsafe control actions were identified in the STPA, which resulted in 8 remaining unsafe control actions after considering the safety barriers. The HAZOP identified 4 hazards, which were equivalent to 9 of the UCAs identified in the STPA, while the FMEA identified 8 failure modes that were equivalent to 14 of the UCAs identified in the STPA. Thus, the HAZOP and the FMEA identified less hazards than the STPA in total, and Leveson's claim about STPA identifying more hazards than traditional hazard identification methods was confirmed (Leveson, 2013). The STPA and the FMEA identified two new hazards that had not been mentioned among the UCAs in the STPA, but the total amount of UCAs identified were greater in the STPA.

Further on, the experiences and evaluation of the STPA workshop that had been conducted, were used together with the suggested STPA framework, theory on STPA and STPA-SEC, and feedback from experts in order to create a final STPA framework. The final design of the STPA framework was inspired by the HAZOP framework and the FMEA framework in "Risk Assessment" (Rausand, 2011), while the actual steps in the analysis were mainly based on "A systems approach to risk analysis of maritime operations" (Rokseth et al., 2017), as well as "STPA Handbook" (Leveson and Thomas, 2018). The framework ended up with 10 main steps, which described a step-by-step approach when conducting an STPA workshop.

There were six main objectives given for the thesis, and all the objectives were met. First, the main system and its relevant functions were described, and all elements in the "secure" function were identified. Both the "secure" function and the main function were illustrated with suitable models. Next, all theory on STPA, STPA-SEC, as well as theory on traditional risk assessment frameworks, were presented and used as a basis for the analysis. The analysis performed in the STPA workshop was described through three main phases, and then compared to the results of the HAZOP and the FMEA workshop. At the end of the report, all the results from the STPA workshop were summarized, and an STPA framework was made.

When considering Bane NOR's requirements to models (Sivertsen, 2017c), the control loop as a model structure does fulfil the requirements, and is therefore suitable as a hazard identification tool for workshops if the participants are familiar with the system. Regarding the structure, the control loop supports the breakdown of the system into constituent parts by using loops, it includes necessary descriptions and system boundaries, it shows the relationship between the components, and it is possible to extend gradually. Further on, the requirements to the behaviour of the system are fulfilled as the control loop shows inputs and outputs, how the system responds to changes, as well as showing actions performed by the system as a whole. The interaction requirements are also fulfilled because the control loop shows how the system can influence or be influenced by the environment, and how the components can be affected by the operation of the system. The control loop shows the relationship between causes, hazards, and accidents, and it did facilitate the identification of all system failure modes that could lead to a dangerous scenario. In addition, it facilitates the identification and treatment of new hazards arising, as well as facilitating the determination of the required safety and all safety related functions. Regarding the design and quality requirements, the design of the control loop made it possible to identify the need for, and analyze the effectiveness of, safety functions or any other barriers, as well as facilitating the demonstration of independence among different functions. The control loop was understandable, included well-defined terms, and it was possible to communicate to the participants in the workshop (Sivertsen, 2017c).

7.2 Discussion

Although the STPA proved to be advantageous and identified more hazards than the HAZOP and the FMEA, there are limitations to the study that might have affected the results and the reliability of the study. There were many differences regarding the execution of the workshop for the three hazard identification methods, which might have affected the results. In the STPA workshop, the safety barriers were taken into account such that the excessive UCAs could be left out. The HAZOP and the FMEA did not identify the remaining hazards after considering safety barriers, and therefore only the UCAs identified before the safety barriers

in the STPA were used for the comparison. Further on, the different methods did not have the same basis for comparison. Both the STPA and the FMEA analyzed the "secure" function only (Gran, 2018), while all the functions were analyzed in the HAZOP (Opsahl, Solibakke and Skogvang, 2018). Therefore, the basis for comparing the STPA and the FMEA were better, and thus gave more accurate answers than the comparison of the STPA and the HAZOP. By only considering one part of the system, it is likely that the results will be more detailed and comprehensive. In addition, the requirements specification that was provided before conducting the STPA might have affected the decision-making process when conducting the STPA (Sivertsen, 2014). It is likely that more UCAs could have been identified by including STPA-SEC in the workshop, which would have been done as explained in Chapter 3, by adding some extra steps concerning security to the original approach (Leveson and Thomas, 2018). By including STPA-SEC, the method becomes more applicable to the increasing number of cybersecurity systems, which is of high importance in today's modern society that is constantly adapting to new technology (Friedberg et al., 2017).

All of the dangers were converted to UCAs in order to make the results measureable, and thus comparable. Because the three hazard identification methods use different terms and definitions for danger, the comparison was not optimal as dangers might have been left out in the workshops. Furthermore, the STPA was performed later in the SafeT project when more knowledge about the system was gained. This could be an advantage and a justification for why the STPA was performed in an effective way, but it might as well be a disadvantage. Some of the participants at the STPA workshop had participated in the HAZOP or the FMEA workshop earlier, and therefore many of the dangers or unsafe control actions were repeated in the STPA. If the participants in a workshop have analyzed the same system earlier, it is likely that it will affect the creative thinking in the team (Rausand, 2011). Nevertheless, representatives from Bane NOR were present during the STPA workshop, and they had technical competence on the detailed responsibilities of each component in the "secure" function, which was a clear advantage and not the case in the FMEA workshop.

STPA is based on a control loop with limited components, while in a HAZOP and an FMEA each function or part of the system is analyzed systematically and based on experience data (Rausand, 2011). The control loop lays the foundation for the whole analysis in an STPA, and therefore the selection of design and details included are important (Rausand, 2011). Even though the control loop fulfils the requirements to models given by Bane NOR (Sivertsen, 2014), such simple models may lose important information and only show direct relationships (Leveson, 2013).

The claim by Leveson that STPA is advantageous to use at complex systems, and that an STPA reveals more hazardous scenarios than the traditional hazard identification methods (Leveson, 2013) is true in some settings, which it is in this thesis where more hazardous scenarios were identified in the STPA. The method is simple and it can be performed in a short period of time, which is advantageous when dealing with a complex system. In addition, it can be performed early in the process and by one person only if necessary (Leveson, 2013).

To summarize, the STPA workshop did provide the results necessary to be able to compare the method to traditional hazard identification methods, as well as creating an STPA framework for practical implementation and facilitation of STPA. More hazards were identified in the STPA than the other hazard identification methods, which confirms Leveson's claim (Leveson, 2013). However, the results of an STPA will depend on different factors like the composition of the STPA team, the amount of preparation beforehand, familiarity with the system, and if the team members have participated in workshops that deal with the same system before (Rausand, 2011).

Regarding the STPA framework, it is inspired by two existing frameworks and methods that are well established today, which should be a good reason for the framework to be valid. On the other hand, it might be that the frameworks used for the HAZOP and the FMEA are not suitable for performing an STPA. The STPA framework design is simple and easy to follow, and it is based on actual experiences and evaluations in addition to literature research. The framework is also based on the actual execution of the workshop, and feedback from experts who participated. These experiences and evaluation process strengthens the reliability of the framework. Nevertheless, the framework is made by one single person, which can make it less reliable and subjective because it is based on personal assumptions and interpretations as well.

Ideally, a group of people that had not participated at any of the other workshops, but are familiar with the system and have knowledge about hazard identification, should have been participants in the STPA workshop. Then, the results of the STPA could have been compared on equal conditions with results from the other workshops that used the SafeT case. Another alternative could have been to choose a different case and perform two workshops (each workshop with a different group of participants), one using the STPA method and another using one of the two other methods for hazard identification. Either of these two alternatives could have given more "unbiased results" to compare STPA with another method.

The railway domain has been in focus when creating the STPA framework in this thesis, but the framework could be applicable to other domains as well where an STPA can be

conducted. Examples of other domains are the automotive domain and the subsea domain (Leveson and Thomas, 2018), as well as the aircraft domain (Leveson and Thomas, 2018). Furthermore, if the STPA is extended to include STPA-SEC, STPA might have other purposes in the development process of a new concept. Then the framework can be applied to more cybersecurity systems, security control, and business and mission analyses in organizations (Young and Porada, 2017).

7.3 Recommendations for Further Work

7.3.1 Short-Term

The comparison between the STPA and the HAZOP and the FMEA, can be developed further by improving the reliability of the results, which can be done by:

- Arranging a workshop with only people who have not worked on the "secure" function before, but still have the required knowledge about the system.
- Either conduct an STPA on the whole system like in the HAZOP, or perform more traditional hazard analyses on the "secure" function only, in order to have a better assessment basis.
- Create a common definition of danger, which covers all of the three definitions in the STPA, the HAZOP, and the FMEA.
- Consider safety barriers in the HAZOP and the FMEA in order to identify remaining hazards, which then can be compared to the remaining UCAs in the STPA.

If these steps are followed, a more accurate comparison could have been done between STPA and the other traditional hazard identification methods, which would result in a more valid conclusion on whether STPA is a suitable hazard identification tool for workshops. In addition, new input might have been added to the system, and new UCAs might have been identified if the participants had not conducted a hazard analysis on the "secure" function previously.

7.3.2 Medium-Term

Further on, the STPA framework for practical implementation and facilitation of STPA can be developed further by:

- Test the STPA methodology in different settings with different participants, and then study the results and feedback given for the different workshops.
- Include the STPA-SEC when conducting the workshop, and then add the new experiences and feedback to the final framework

By following these steps, the STPA framework can be improved and developed further. In addition, if the implementation of STPA-SEC is based on experiences and feedback from the workshop as well, an even more realistic framework can be created.

7.3.3 Long-Term

The STPA can also be expanded and developed further by:

- Insert the UCAs into a risk model where they are linked to hazards and barriers to see which barriers that already exist in the existing concepts, and which barriers that could be appropriate to introduce to the original concept
- Study the RAM aspects of the system simultaneously when performing the analysis
- Add automated tools and logic tables to the practical implementation of STPA

If these steps are followed, the requirements specification given by Bane NOR could have been improved by adding extra requirements to the system, and modifying already existing requirements. Furthermore, the analysis would not be limited to safety only if the RAM aspects were studied simultaneously. In addition, implementation of automated tools and logic tables could have simplified the facilitation and recording of the STPA workshop, as well as reducing the time spent on it.

Bibliography

Dong, A. (2012). Application of Cast and STPA to Railroad Safety in China. [ebook] Massachusetts, pp.3,14-81. Available at: http://sunnyday.mit.edu/safer-world/Airong-thesis.pdf [Accessed 5 Jan. 2018].

Eltervåg, A., Hansen, T., Lootz, E., Rasmussen, E., Sørensen, E., Johnsen, B., Heggland, J., Lauridsen, Ø. and Ersdal, G. (2017). Barrierenotat 2017. 3rd ed. [ebook] Petroleumstilsynet, pp.3-5. Available at: http://www.ptil.no/getfile.php/1343444/PDF/BARRIEREnotat%20%202017.pdf [Accessed 5 Jun. 2018].

Friedberg, I., McLaughlin, K., Smith, P., Laverty, D. and Sezer, S. (2017). STPA-SafeSec: Safety and security analysis for cyber-physical systems. Journal of Information Security and Applications, [online] 34, pp.183-196. Available at: https://ac.els-cdn.com/S2214212616300850/1-s2.0-S2214212616300850-main.pdf?_tid=80679ac7-bedb-42ed-9a64-828a68b535ed&acdnat=1529593185_876483ef1a235f2a1e0ce9c153ad7d5f.

Gran, B. (2018). FMEA function secure. Halden, pp.1-5.

Gran, B., Karpati, P. and Hauge, A. (2018). SafeT- Developing the Risk Approach. pp.1-10.

Hansen, S. (2018). SafeT- Next Generation Safety Assessment Framework for Railway. Trondheim, pp.15-30.

Jernbaneverket (2015). ERTMS Implementation Plan. [online] pp.2-3. Available at: http://www.banenor.no/globalassets/documents/ertms/ertms-implementation-plan-eng.pdf [Accessed 21 Jun. 2018].

Jernbaneverket (2013). Sikkerhetshåndboken- Sikkerhetsstyring i Jernbaneverket. [online] p.32. Available at: http://www.banenor.no/contentassets/65c101dfc8304e93b8c4e99a96ff3928/sikkerhetshandboken-10-7-2013.pdf [Accessed 18 Jun. 2018].

Leveson, N. (2013). An STPA Primer. 1st ed. [ebook] MIT, pp.3-72. Available at: http://sunnyday.mit.edu/STPA-Primer-v0.pdf [Accessed 4 Jan. 2018].

Leveson, N., Young, W. (2014). STPA-SEC for Cyber Security / Mission Assurance. [ebook] pp.3,5-12,19. Available at: http://psas.scripts.mit.edu/home/wp-content/uploads/2014/03/Young_STAMP_2014_As-delivered.pdf [Accessed 7 Jun. 2018].

Leveson, N., Young, W. (2013). Systems Thinking for Safety and Security. [ebook] Cambridge, pp.1-8. Available at: http://delivery.acm.org/10.1145/2540000/2530277/p1-young.pdf?ip=129.241.229.203&id=2530277&acc=ACTIVE%20SERVICE&key=CDADA77FFDD8BE08%2E5386D6A7D247483C%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1529002944_b0a28c55611f3d3e118dab3f357330fb [Accessed 14 Jun. 2018].

Opsahl, R., Solibakke, S. and Skogvang, Ø. (2018). SafeT- Hazard Identification. [online] Trondheim, pp.8-11,17,20. Available at: http://www.bim2share.no/download/53616c7465645f5f632813c7b03e273ae7a81ff9f6e1ce5ca5836fff9163989ce4d57baa0a16a398/SafeT%20-%20HAZOP%20report%20%28textual%20description%29.pdf [Accessed 16 Jun. 2018].

Rausand, M. (2011). Risk assessment. Hoboken, N.J: Wiley, pp.236-256.

Rokseth, B., Bouwer Utne, I. and Vinnem, J. (2017). A systems approach to risk analysis of maritime operations. [ebook] SAGE Publications, pp.7-30. Available at: https://brage.bibsys.no/xmlui/bitstream/handle/11250/2464837/Postprint%20-%20A%20systems%20approach%20to%20risk%20analysis%20of%20maritime%20operations.pdf?sequence=4&isAllowed=y [Accessed 4 Jan. 2018].

Sivertsen, T. 2017a. Concept of a New Solution for Securing Work Areas. Oslo, pp.3-22.

Sivertsen, T. 2017b. SafeT – Case example on securing work areas. Oslo, pp.1-14.

Sivertsen, T. 2017c. SafeT- Requirements to models. Oslo, pp.2-9.

Sivertsen, T. (2014). Kravspesifikasjon for ny løsning for sikring av arbeid i og ved spor. pp.92-117.

Smith, D. (2011). Reliability, Maintainability and Risk. 8th ed. Burlington: Elsevier Science, pp.137-141,164-169.

Stangeland, H., Lundon, I. and Skogvang, Ø. (2018). Workshop Risikovurdering IKDP Forus, pp.1-10

The what, why and how of the GSM-R System. (2017). [ebook] NetworkRail, pp.2-3. Available at: https://www.networkrail.co.uk/running-the-railway/gsm-r/ [Accessed 5 Jan. 2018].

Young, W. and Porada, R. (2017). System-Theoretic Process Analysis for Security (STPA-SEC): Cyber Security and STPA. [ebook] Boston, pp.24-66. Available at: http://psas.scripts.mit.edu/home/wp-content/uploads/2017/04/STAMP_2017_STPA_SEC_TUTORIAL_as-presented.pdf [Accessed 12 Jun. 2018].

Appendix A: Acronyms

APP Application

CA Control action

CL Control loop

CPS Cyber-physical system

CTC Centralized traffic control

ERTMS European railway traffic management system

FMEA Failure modes and effects analysis

FMECA Failure mode, effects and criticality analysis

FTA Fault tree analysis

GSM-R Global system for mobile communications-railway

HAZOP Hazard and operability study

ICT Information and communication technology

IFE Institute for energy technology

PM Process model

PMV Process model variable

QR Quick response

RAMS Reliability, availability, maintainability and safety

SC Safety constraint

SG Safety guard

SLA System-level accident

SLC System-level constraint

SLH System-level hazard

STAMP Systems- Theoretic Accident Model and Processes

STPA System theoretic process analysis

STPA-SEC System theoretic process analysis for security

SUS Support system

SYSML Systems modelling language

TD Train dispatcher

TFFR Tolerable functional failure rate

THR Tolerable hazard rate

UCA Unsafe control action

WA Work area

Appendix B: Bane NOR's Requirements to Models

The safety requirements are based on the CENELEC standards for functional safety, which are EN 50126, EN 50128, and EN 50129. The requirements on models given by Bane NOR are (Sivertsen, 2017c):

Structure

The models must:

- Support the breakdown of a system into its constituent parts, in terms of system, sub-systems, and components
- Support any hierarchy of system levels, and the possibility to describe any system level at the appropriate level of detail without introducing unnecessary detail and complexity at other system levels
- Support precise descriptions of the physical and functional boundaries of each element
- Facilitate the treatment of systems, sub-systems and components as black boxes, for which the details on architecture, design and implementation can be kept out of consideration, evaluating functions and hazards only at the boundaries
- Show the relationship and intended integration between hardware and software
- Be able to show the functions and structure of hardware and software to any level of detail
- Be possible to extend gradually, supporting activities related to concept, via risk assessment and hazard control, to safety demonstration and assessment

Behaviour

The models must:

- Show how the behaviour and state of a system depends on, and changes with, the functionality of its sub-systems and components
- Show how a system responds to inputs to produce specified outputs, whilst interacting with its environment
- Show the actions or activities performed by the system as a whole, and distinguish these from those internal to the system
- Discern between function and structure, but show how these are related to each other

Interaction

The models must:

- Describe the system as contained in its operational environment

- Show how the environment can influence, or be influenced by, the system, including anything to which the system connects mechanically, electrically or by other means
- Show how man and organization can affect, or be affected by, the operation of the system
- Facilitate the derivation of the system's possible impact on the RAMS performance of neighbouring systems

Risk

The models must:

- Facilitate the identification of hazards associated with the system and events leading to these hazards, the determination of the risk associated with the hazards, and the identification of possible further safety requirements needed to reduce the risk to an acceptable level, at any system level
- Facilitate the identification of all system failure modes that could lead to a hazard, and the analysis of the consequences these failure modes can have on the functionality of the system
- Facilitate the identification of all sequences and/or coincidences of events, failures, operational states, environmental conditions, etc. that could result in an accident
- Show the relationships between causes, hazards, and accidents
- Make it possible to determine if a hazard can be practicably avoided, or otherwise how the risk can be reduced
- Facilitate the identification of the degree of safety required for each particular situation
- Facilitate the identification and treatment of new hazards arising from design
- Facilitate refinement of the hazard identification as new system details and levels are introduced
- Facilitate the determination of all safety related functions
- Facilitate the determination of the required safety integrity of all safety related functions
- Support techniques for hazard identification, consequence analysis, and common cause failure analysis
- Facilitate the identification of physical and information technology security threats that might have an impact on functional safety

Requirements

The models must:

- Provide the details necessary to explain and understand the requirements to the functions to be provided by the system, as well as any additional requirements that are necessary to ensure proper functioning, including contextual and technical requirements
- Facilitate demonstration of the completeness of the safety requirements
- Facilitate apportionment of the RAMS requirements to the different subsystems and components, based on a precise definition of the process boundaries and boundary conditions
- Facilitate the derivation of tolerable hazards rate (THR) and, where relevant, tolerable unsafe functional failure rates (TFFR), at any system level
- Support the systematic breakdown of safety targets (THR/TFFR), as appropriate for the allocation of safety integrity levels (SIL)
- Facilitate the identification of safety requirements
- Be traceable to/from the safety requirements

Design

The models must:

- Be possible to refine into design descriptions that verifiable implement the functional requirements
- Make it possible to determine how a function can be made fail-safe, if at all possible
- Make it possible to identify the need for, and analyze the effectiveness of, safety functions or any other barrier
- Facilitate the analysis of functional independence and common cause failures
- Support the explanation of the technical principles which assure the safety of the design
- Facilitate the demonstration of the effectiveness of safety functions and barriers
- Facilitate the analysis of the effects of faults as appropriate for the safety demonstration
- Facilitate the demonstration of independence among functions
- Facilitate the demonstration of independence from common random causes
- Facilitate the demonstration of independence from common systematic causes

Quality

The models must:

- Use clear and intelligible means of description, such as formal notation for logical functions, natural language for introductions, justifications and representations of

intentions, graphical representations of examples, semantic definition of graphical elements, and directories of specialized words

- Be adequate to the requirements engineering process, in particular to the elicitation and analysis of safety requirements
- Have well-defined syntax and semantics, and otherwise suitable for tools support
- Be possible to communicate to the different stakeholders
- Be possible to review for completeness of the identified safety requirements
- Prevent inconsistent descriptions
- Be possible to assess with respect to the necessity and sufficiency of the level of detail
- Be understandable in themselves
- Be understandable to the prospective user
- Be unambiguous
- Use adequate notation
- Use well-defined terms and graphical element

Appendix C: Additional Tables Used for the STPA Analysis

Appendix C contains additional tables that were not included in the STPA analysis, as well as complete tables for the tables where only a small part was included.

Controller Responsibilities and Process Models

Role	Responsibilities	Input/process model (PM)
Train dispatcher (TD)	***Loop 1:*** ●Receive call from SG ***Loop 4:*** ●Block work area, and release it for securing	***Loop 1:*** ●Block request (block/unblock) ***Loop 4:*** ●Status WA (secured/released)
Safety guard (SG)	***Loop 1:*** ●Call TD ***Loop 2:*** ●Scan work area (WA) ●Secure WA ●Confirm secured WA	***Loop 2:*** ●Secure request (accepted/rejected)
App	***Loop 2:*** ●Ask SG to scan WA ***Loop 6:*** ●Send request to SuS to secure WA ●Send message to SuS with confirmed WA	***Loop 2:*** ●Scan WA (Identified/not identified) ●Secure WA (Secure/release) ●Confirm WA (confirmed/not confirmed) ***Loop 3:*** ●SG position (on WA/outside WA) ●Status WA (identified/not identified) ***Loop 6:*** ●Status WA (secured/released) ●Secure permission (granted/declined) ●Secure confirmation (secured/not secured) ●Alarm, wrong WA (emergency/no emergency)
QR code	***Loop 3:*** ●Identify WA ●Identify SG's position ●Send confirmation to SG	***Loop 3:*** ●Scan QR (scanned/not scanned)

Support system (SuS)	***Loop 5:*** ●Check if SG can secure WA (if not already secured) ●Check if WA is blocked and released for securing ***Loop 6:*** ●Accept/reject request from App to secure WA ●Request SG to confirm WA ●Check that confirmed WA = requested WA ●Associate SG to WA	***Loop 5:*** ●Status WA (secured/not secured) ***Loop 6:*** ●Smartphone (secure/release/ scan) ●Status WA (confirmed/not confirmed)
Centralized traffic control (CTC)	***Loop 5:*** ●Give clearance to unblock WA when securing is released	***Loop 4:*** ●Status WA (blocked/unblocked) ***Loop 5:*** ●Status WA (secured/released/blocked/unblocked)

Observation Form

Role	Corporation/ Expert	Responsibility (control action)	Scenarios	Input	UCA	Mitigation/ safety barrier	Remaining UCA
SG/TD	Bane NOR (technical experts): BN Safetec (experts on hazard identification and risk analysis): ST IFE (experts on STPA and hazard identification): IFE	SG: Call TD Secure WA Scan WA code Confirm selected WA TD: Receive call Block WA, and release for securing	SG does not call TD SG calls, TD does not respond Call results in unsafe situation SG calls too early SG calls too late SG hangs up too soon TD responds too early	Block status (from TD) Status WA (from App)	SG hangs up too early, and does not get confirmation of block status (BN) SG makes the call before arriving at the WA (IFE) TD is not available when SG calls (IFE) SG secures a small WA in a larger on. SG joins a secure WA (BN) SG does a full securing without	Support system checks status of blocking and confirms release (BN) All functions will be stopped when previous conditions are not met (BN) SuS will be notified TD is not available(BN) SG must scan out of the WA if he has two WAs All functions will be stopped when previous conditions are not met. (BN)	SG makes the call before arriving at the WA SG secures the WA before calling TD TD unavailable when there is a change of guards

			TD responds too late TD hangs up to soon TD hangs up too late	Secure status (from SG)	involving SG (IFE) TD blocks wrong WA with another intention (ST) Not able to call TD because of system down (ST) TD does not keep line until WA is secured (BN) Status incorrect because of bad reception (IFE) App does not work (ST) Vandalism such that code is not available (BN)	Procedure should include questions to ask SG during securing (BN) GSM-R phone can be used instead, and SuS and SMS can be used for communication. Procedure for how long to hold the line (ST) GSM-R phones used (BN) SuS notified (BN) Extra copies must exist. GPS can be used (ST)	TD does not confirm that the WA is blocked
App	Bane NOR (technical experts): BN Safetec (experts on hazard identification and risk analysis): ST IFE (experts on STPA and hazard identification): IFE	Ask SG to scan WA Send req. to SuS to secure WA Send message to SuS with confirmed WA	App does not ask SG to scan App ask SG to scan too late App asks to scan wrong WA Req. to SuS not sent	Status WA (from SuS) Command SG (SG)	SG calls too late, move to WA before it is safe (IFE) Both SMS and App system is down (BN) Not secured, secured too late, secured too early (IFE) Function failure in the	All moving trains will stop (BN) All functions will be stopped when previous conditions are not met (BN) Data should be cleared from	SG calls too late, move to WA before it is safe

			Req. to SuS sent too late/too early Message to SuS not sent Message to SuS sent too late/too early		system-cannot stop it if scanned WAs do not correlate (ST) SG pushes the wrong button (ST) SG scans two times in a row, confirmation too early (ST) SG send legal, but not valid code. Interpreted as wrong WA (IFE) Confirmed WA has already been covered (ST)	data system regularly (ST) SuS supervises this protocol (BN) Dual coding (ST)	SG pushes the wrong button SG send legal, but not valid code. Interpreted as wrong WA
CTC	Bane NOR (technical experts): BN Safetec (experts on hazard identification and risk analysis): ST IFE (experts on STPA and hazard identification): IFE	Give clearance to unblock WA when securing released	Does not give clearance Give clearance too late Give clearance too early Gives wrong clearance	Status WA (from SuS) Block status (from TD)	Wrongly programmed CTC (IFE) Shows clear when WA is blocked (IFE)	Extra barriers on the supervision (BN) It must be checked with SuS before clearance to unblock is given (BN)	
SuS	Bane NOR (technical experts): BN Safetec (experts on hazard identification and risk analysis): ST	Check if SG can secure WA Check if WA is blocked & released for securing Ensure WA cannot be unblocked before			SuS accept/does not accept the securing (IFE)	CTC supervises the SuS actions (BN)	

	IFE (experts on STPA and hazard identification): IFE	securing released Accept/reject request					

UCAs, Safety Barriers and Remaining UCAs

ID	UCA	Already existing safety barrier	Planned safety barrier, SC (from Bane NOR/experts)	Remaining UCA
1	SG does not call TD	TD is in a "deadlock", and no securing or blocking is possible until the call is made		
2	SG makes the call before arriving at the WA			1. SG makes the call before arriving at the WA
3	SG loses data connection when calling TD		It is not possible to continue to the next step in the "secure" process until the connection is re-established	
4	SG hangs up too early and does not get confirmation on block status	All functions will be stopped when previous conditions are not met.		
5	SG is not able to secure WA since WA is already secured by others	The App will notify the SG if the App receives message from the SuS that the WA is already secured		
6	SG secures WA too early	All functions will be stopped when previous conditions are not met.		
7	SG secures WA too late, and move to WA before it is safe	Covered by existing operating rules		

8	A full securing is performed without involving TD at all	All functions will be stopped when previous conditions are not met.		
9	SG pushes the wrong button on the App when securing			2. SG pushes the wrong button on the App when securing
10	SG secures WA before calling TD	All functions will be stopped when previous conditions are not met.		
11	SG secures a smaller WA in a larger one, gets message that WA is already secured		SG must scan out of the WA if in charge of two work areas at the same time	
12	SG does not scan the QR code	It is not possible to continue to the next step in the “secure” process until the code has been scanned		
13	SG scans the wrong QR code		A GPS system will check that the SG is in the right area when scanning work area	
14	SG scans the QR code two times in a row, such that the confirmation is given too early			3. SG scans the QR code two times in a row, such that the confirmation is given too early
15	TD is not available when SG calls	Covered by existing operating rules		
16	TD does not respond, and SG secures WA anyways	TD is in a “deadlock”, and no securing or blocking is possible until the call is made		
17	Data connection is lost, but TD still has voice connection	If the GSM-R system is down, no trains will be running and all maintenance work will be stopped.		

18	TD forgets to block WA, and confirm it as blocked to SG	WA must be blocked before it can be secured, and SuS checks that the actual WA is ready for securing.		
19	TD blocks the WA too late such that trains can enter a secured WA	SuS will check with CTC that WA suggested blocked is actually blocked and released for securing		
20	TD blocks the WA too early, before the WA is secured	All functions will be stopped when previous conditions are not met.		
21	The App does not ask the SG to scan the relevant WA	All functions will be stopped when previous conditions are not met.		
22	The request from the App to the SG is not sent	SuS is supervising this action, and the “secure” process will not continue until the request is sent		
23	TD blocks the WA, but the “secure” function on the App does not work	Covered by existing operating rules		
24	A legitimate but incorrect code is sent, and wrong WA is confirmed	SuS is supervising this action, and makes sure that confirmed WA=requested WA		
25	The status shown from the App is incorrect			4. Wrong information on the status of the WA is given to the SG 5. The app distorts the information from SuS, even though it is transferred correctly
26	WA requested has been covered earlier, and no confirmation is given			6. WA requested has been covered earlier, and no confirmation is given

27	The QR code has been counterfeited such that it is not possible to identify the WA	TD shall immediately report if there is a suspicion of mismatch between the code sign at the WA and the code		
28	It is not possible to scan the QR code because of vandalism	The QR code is only determined by the information on the sign, and easy to reproduce. The production of the sign should be automated as much as possible		
29	The SG's position is identified before the WA has been identified		The WA will be identified first when the SG scans the WA, and a GPS system will check that the SG's position is correct	
30	A copy of the QR code has been made such that the code can be scanned without the SG physically being at the WA		A GPS system will be used to the check the SGs' positions to make sure that the positions on the App are correct	
31	A confirmation is not sent to the SG because the App system is down	If the GSM-R system is down, no trains will be running and all maintenance work will be stopped.		
32	Confirmation is sent too early to the SG such that securing is started before the WA is confirmed ready	All functions will be stopped when previous conditions are not met.		
33	CTC gives clearance to unblock WA before securing is released	SuS must confirm that the WA can be unblocked before clearance is given		
34	Clearance from the CTC is incorrect, and a secured and unreleased WA is unblocked	SG and/or TD will physically check the WA before unblocking a WA.		

35	Clearance from the CTC is incorrect, and the wrong WA is unblocked	SuS must confirm that confirmed WA= requested WA		
36	Wrong status on the WA is given because of malfunction in the system	SuS must confirm that confirmed WA= requested WA		
37	A secure request is rejected when the WA is supposed to be secured	SuS supervises the CTC's actions		
38	A secure request is accepted when the WA is not supposed to be secured	SuS supervises the CTC's actions		
39	A confirmation that the WA can be secured/released is not sent from the SuS to the App	All functions will be stopped when previous conditions are not met.		
40	A confirmation that the WA can be secured/released is sent too early from the SuS to the App	All functions will be stopped when previous conditions are not met.		
41	A confirmation that the WA can be secured/released is sent too late from the SuS to the App			7. A confirmation that the WA can be secured/ released is sent too late from the SuS to the App
42	SG is not associated to a WA such that it is unknown whether some of the SG's are still at the WA when it is released		A GPS system will check if there are any SGs left on the WA before the WA is released.	
43	SuS associates SG to WA, but SG switches WA without registering in the App			8. SuS associates SG to WA, but SG switches WA without registering in the App
44	SuS associates SG to wrong WA		A GPS system will check that the SG is at the right WA	

Dangerous Scenario, UCA, and Associated Causal Factor

Role	Control action/ responsibility	Mode/scenarios	UCA	Causal factor
SG	Call TD	SG does not call the TD	1. SG forgets to call TD	Human error, SG has an inattentive moment
		SG calls the SG too early	2. SG makes the call before arriving at the WA	Human error, the SG is not supposed to call before entering the WA
		SG keeps the line too short	3. SG loses data connection when calling TD	GSM-R system is down/system failure
			4. SG hangs up too early and does not get confirmation on block status	Human error, SG should wait for the confirmation before hanging up
	Secure WA	SG does not secure the WA	5. SG is not able to secure WA since WA is blocked for any actions	System does not allow the SG to secure a WA that is already in use
		SG secures the WA too early	6. SG secures WA too early	SG does not wait for the confirmation from TD before securing the WA
		SG secures the WA too late	7. SG secures WA too late, and move to WA before it is safe	SG does not secure the WA right after the confirmation from the TD is given, and does not check that the WA is safe before entering.
		Securing results in an unsafe situation	8. A full securing is performed without involving TD at all	It is SG's responsibility to involve TD by making a call before securing.
			9. SG pushes the wrong button on the App when securing	Human error, SG is not observant when using the

				App for securing the WA.
		SG secures the WA at the wrong time	10. SG secures WA before calling TD	Human error, SG does not get a confirmation from TD before securing the WA.
			11. SG secures a smaller WA in a larger one, gets message that WA is already secured	System denies SG to secure two WA's, even though one WA is a part of the other WA.
	Scan QR code	SG does not scan the QR code	12. SG forgets to scan the QR code	Human error, SG has an inattentive moment
		Scanning results in an unsafe situation	13. SG scans the wrong QR code	Either SG has an inattentive moment, or the QR code is placed at the wrong WA.
	Confirm selected WA	SG confirms the WA at the wrong time	14. SG scans the QR code two times in a row, such that the confirmation is given too early	Either SG does not wait for the App asking the SG to scan the WA before scanning it or there is an error in the App such that the request to scan never appears to the SG.
TD	Receive call	TD does not receive the call form the SG	15. TD is not available when SG calls (changing guards)	Human error, the change of TDs is not overlapping.
			16. TD does not respond, and SG secures WA anyways	SG does not wait for the confirmation from the TD before securing the WA.
		TD keeps the line too short	17. Data connection is lost, but TD still has voice connection	System failure or the system is down, but the GSM-R network still works.

	Block WA, and release for securing	TD does not block the WA	18. TD forgets to block WA, and confirm it as blocked to SG	Human error, TD has an inattentive moment.
		TD blocks the WA too late	19. TD blocks the WA too late such that trains can enter a secure WA	Human error, TD does not block the WA right after receiving the call from the SG. Also system error as it should not be possible for trains to enter before the WA is blocked.
		TD blocks the WA too early	20. TD blocks the WA too early, before the WA is secured	Human error, TD does not wait for confirmation on secured WA from the SG before blocking. Also system error as it should not be possible to block a WA that is not secured.
		Blocking results in an unsafe situation	21. TD blocks another WA with another intention, and secure function goes through	Human error, TD has an inattentive moment and blocks the wrong WA.
App	Ask SG to scan WA	Request from the App to scan the WA is not sent	22. The App does not ask the SG to scan the relevant WA	plication error
	Send request to secure WA	Request to secure WA is not sent	23. The request from the App to the SG is not sent	plication error
		Sending the request to secure WA results in an unsafe situation	24. TD blocks the WA, but the “secure” function on the App does not work	Application error or smartphone error.
	Send message to SuS with confirmed WA	Message sent to SuS with confirmed WA results in an unsafe situation	25. A legitimate but incorrect code is sent, and wrong WA is confirmed	System failure in the SuS, SuS is not able to confirm that requested WA=confirmed WA.
			26. The status shown from the App is incorrect	Application error or failure in the SuS.

		Message to SuS with confirmed WA is sent at the wrong time	27. WA requested has been covered earlier, and no confirmation is given	App is not able to cover two WAs with different intentions .at the same time.
QR code	Identify WA	QR code does not identify the WA	28. The QR code has been counterfeited such that it is not possible to identify the WA	Somebody is intending to hack the App.
			29. It is not possible to scan the QR code because of vandalism	The QR code sign has been ruined, and is not readable.
		QR code identifies the WA at the wrong time	30. The SG's position is identified before the WA has been identified	The App has not requested the SG to scan the WA before locating the SG.
	Identify SG's position	Identifying the SG's position results in an unsafe situation	31. A copy of the QR code has been made such that the code can be scanned without the SG physically being at the WA	SG wants to simplify the task such that it is not necessary to be at the WA.
	Send confirmation to SG	Confirmation is not sent to SG	32. A confirmation is not sent to the SG because the App system is down	Either there is an Application error or a system failure in the SuS.
		Confirmation is sent to SG too early	33. Confirmation is sent too early to the SG such that securing is started before the WA is confirmed ready	Either there is an Application error or a system failure in the SuS. Securing is started without permission from SuS/App.
CTC	Give clearance to unblock WA when securing is released	Clearance to unblock the WA is given too early	34. CTC gives clearance to unblock WA before securing is released	Error in the CTC system results in an incorrect clearance.
		The clearance given to unblock the WA results in an unsafe situation	35. Clearance from the CTC is incorrect, and a secured and unreleased WA is unblocked	Error in the CTC system results in an incorrect clearance.

			36. Clearance from the CTC is incorrect, and the wrong WA is unblocked	Error in the CTC system results in an incorrect clearance.
SuS	Check if SG can secure WA	Checking if the SG can secure the WA results in an unsafe situation	37. Wrong status on the WA is given.	Malfunction in the system or in the Application.
	Check if WA is blocked and released for securing		37	
	Ensure WA cannot be unblocked before securing is released		37	
	Accept/reject secure request	SuS does not accept/reject secure request	38. A secure request is rejected when the WA is supposed to be secured	System failure in the SuS leads to incorrect rejecting of a request.
			39. A secure request is accepted when the WA is not supposed to be secured	System failure in the SuS leads to incorrect accepting of a request.
	Send confirmation to App	SuS does not send a confirmation to the App	40. A confirmation that the WA can be secured/released is not sent from the SuS to the App	System failure in the SuS.
		SuS sends a confirmation to the App too early	41. A confirmation that the WA can be secured/released is sent too early from the SuS to the App	System failure in the SuS.
		SuS sends a confirmation to the App too late	42. A confirmation that the WA can be secured/released is sent too late from the SuS to the App	Either system failure in the SuS or delays in the data connection.
	Request SG to confirm WA		32, 33	
	Check that confirmed		37	

	WA= requested WA			
	Associate SG to WA	SuS does not associate the SG to the WA	43. SG is not associated to a WA such that it is unknown whether some of the SGs are still at the WA when it is released	Either a human error where SG has not reported the position to the SuS via the App, or application error/system malfunction.
		SuS associates the SG to the WA, but the instructions are not followed	44. SuS associates SG to WA, but SG switches WA without registering in the App	Human error, SG does not follow the routine to register in the App when leaving or entering a WA.
		Associating the SG to the WA results in an unsafe situation	45. SuS associates SG to wrong WA	stem failure in the SuS.

Evaluation and Comments to the Workshop

	Planning/preparation	**Execution of the workshop**
Direct feedback	***IFE:*** • A good idea to send out the presentation used in the workshop in advance • The control loop gave a good overview of the system • The use of different colours in the control loop makes it easier to differ between the different functions	***IFE:*** • The method used in the workshop must be based on who the participants are, and what type of knowledge they hold • Post work: Customize the control loop to make it relevant for the group to present it to stakeholders • Positive that the facilitator does not strictly follow the scheme, and opens up for discussion throughout the workshop

	Bane NOR: • Transparent control loop, actions and time perspective on the loops can be included • An STPA must be customized to the group that is present in the workshop	***Bane NOR:*** • Can be confusing with both a control loop and a figure of the function • The control loop is best suited for people familiar to the system since it is advanced • A HAZOP workshop probably will find many of the same hazards • There are a lot of potential in the STPA methodology, it opens up for more creative thinking and does not follow a strict structure • STPA does not reflect the physical part of the system in a good way
	Safetec: • Advantageous with a control structure that shows the interconnection, makes it possible to identify other hazards • Advantageous to print figures to be used during the analysis in advance to all the participants	***Safetec:*** • May be a disadvantage that all the participants are familiar and have a lot of knowledge about the system already • This approach to STPA can be useful and open up to innovative brainstorming • Advantageous that the STPA is performed in a short amount of time • Good discussions are started right away when there is a high level of knowledge about the system among the participants

Feedback from the survey (anonymous)	• Good presentation distributed prior to the workshop • The information given about STPA and the SafeT case in the presentation was very good • Useful control loop for people familiar with the system • Valuable figure of the "secure" function, and description of the steps	• The graphical diagram showing the Loop approach of the STPA method applied to the SafeT case was effective in presenting the starting point as well as providing a tool for "carrying out" the workshop using the STPA method. • The STPA approach worked well because of the participants in the workshop. The approach may not be ideal for people introduced to the system for the first time. • The STPA method was effective to identify the actors (e.g. App, Support system, etc.) and interactions between pairs of actors. • In a process that involves more than two actors, it seems difficult to follow the whole flow of the interactions. • Ideally, a new group of people that had not been at any of the other workshops should have been the active participants in the STPA workshop.

Appendix D: Survey- Evaluation of the STPA Workshop

Appendix D contains results from the survey that was sent out to all the participants after the STPA workshop.

Rating Scale

Rating	Description
1	Very poor
2	Poor
3	Fair
4	Good
5	Very good

Result of the Survey

Number	Question	Answer				
1	How was the information sent out in advance?	4	4	5	5	5
2	How well did the number of participants work?	4	4	5	5	5
3	How was the information given about the system and the STPA at the presentation?	4	4	5	5	5
4	How useful was the control loop given at the workshop?	3	3	3	4	4
5	How was the time frame given for the STPA (analysis)	4	4	4	4	5
6	How was the agenda for the workshop?	4	4	5	5	5
7	How well were the objectives for the STPA achieved?	3	3	4	4	5
8	How well were the objectives for the workshop achieved?	3	4	4	5	5
9	How well-organized was the workshop?	4	4	5	5	5
10		Comment field only				

Additional Comments

Number	Question	Comment
4	How useful was the control loop given at the workshop?	The control loop was useful, and I would rate it between 4-5. However, if I wasn't familiar with the system beforehand, then I might rate it lower (between 3-4).
		Equally valuable was the other complementing drawings and descriptions. The set of descriptions is more valuable than any single description.
10	What worked well? What did not work well? What could have been done differently?	It is hard to rate the STPA approach based on the system considered in the workshop. As my answers reflect, the approached worked well much because of the people in the meeting. Using this approach in a work meeting where people are for the first time

		introduced to the system, might not result in ideal results.
		What worked well? The PowerPoint presentation distributed prior to the workshop and the information given about the STPA method and SafeT case at the STPA Workshop were very good. The graphical diagram showing the Loop approach of the STPA method applied to the SafeT case was effective in presenting the starting point as well as providing a tool for "carrying out" the workshop using the STPA method. What did not work well? The STPA method was very effective to identify the actors (e.g. App, Support system, etc.) and interactions between pairs of actors. However, in a scenario of a process that involves more than two actors, it seemed difficult to follow the whole flow of the interactions. That could be a limitation of the ability of the method to allow discovery of hazards in complex processes with many actors. With the "biased starting point" of having identified hazards in the case using a different method previously, it is difficult to be able to access the effectiveness of using the STPA method to identify hazards. What could have been done differently? The STPA Workshop had a different context or starting conditions compared to the workshops held previously in Halden and Oslo. In those two workshops, the SafeT case was not analyzed by the participants previously with respect to performing a hazard search/discovery. In the STPA workshop, most if not all of the participants had performed a hazard search on the SafeT case previously using another method. Of course, there are limits on time and participants in the SafeT project and that is most likely the reason for choosing SafeT again as the case for the purpose of comparison of results with other methods. Ideally, a new group of people that had not been at any of the other workshops should have been the active participants in the STPA workshop. Then the results of the hazard search/discovery process could have been compared on equal conditions with results from the other workshops that used the SafeT case. Another alternative could have been to choose a different case and perform two workshops (each workshop with a different group of participants), one using the STPA method and another using one of the two other methods for hazard search/discovery. Either of these two alternatives could have given more "unbiased results" to compare STPA with another method.
		Everything worked well. As a follow-up, it would be interesting to see what are the main difference between e.g. STAMP and HAZOP if these are strictly applied according to how they are described and what is the main difference when applying variants, e.g. the variant used in the workshop with commonly applied variants of HAZOP. Are they very similar or not?