A Harmonized Threat/Hazard Modeling Method for Safety Critical Industrial Systems

Andreas Platschek
Institute of Computer Technology, Vienna University of Technology
Gußhausstraße 27-29/384, 1040 Vienna, Austria
platschek@ict.tuwien.ac.at

Abstract

Since the current common practice is to connect every industrial system to the internet in one way or the other, the security of a system has to be evaluated and assured - especially when it comes to safety critical systems.

Recent standards (notably IEC 61508 Ed2 and EN 50159 Ed2), have begun to normatively include security for systems that are no longer closed. These standards contain clauses that require a systematic method used to perform a threat analysis if they could constitute a relevant safety impact. While there is a number of threat modeling techniques available, many of those were developed for the server and office space, but would require a number of adaptions for the use in industrial systems. Other methods are newly developed for industrial systems, but they lack the confidence a development team has to put into them.

A third option - presented in this paper - is to reuse a method that has already been in use in the safety domain for a long time, is well known, understood and trusted, and adapt it to be suitable for security. The methods are compliant with the safety standards and thus the extension - if done carefully - does not invalidate this acceptance and can build on well established competence of the safety engineering staff. At the same time, this harmonization is crucial as both security and safety are system properties and treating interdependent system properties as independent is technically not reasonable and economically not efficient.

The advantage of this approach is, that the development team only needs to be firm in one analysis method and use it for threat analysis when security is analyzed and hazards when safety is analyzed.

1 Introduction

In recent years, the Internet has not only conquered the office and server space as well as our private lives, but also industrial systems. This is of course also reflected in the buzz-words swirling around, like the "Internet of Things (IoT)", "Smart Factory" or the "Industry 4.0". All these buzz words include (amongst other things) the connection of the system to an open network – most of the time the Internet, putting a security context on all of those industrial systems.

All of this becomes even more concerning, when the industrial systems in questions are safety critical systems, as this introduces a situation where the safety of the system depends on the security of the system. In response to this trend of connecting industrial, safety critical systems to open networks, recent versions of safety standards have adapted clauses on security issues.

In example, IEC 61508-1 Ed2, clause 7.4.2.3 states:

The hazards, hazardous events and hazardous situations of the EUC and the EUC control system shall be determined under all reasonably foreseeable circumstances (including fault conditions, reasonably foreseeable misuse and malevolent or unauthorised action). This shall include all relevant human factor issues, and shall give particular attention to abnormal or infrequent modes of operation of the EUC. If the hazard analysis identifies that malevolent or unauthorised action, constituting a security threat, as being reasonably foreseeable, then a security threats analysis should be carried out. [IEC 61508-1, 7.4.2.3]

In addition to this requirement of conducting a security threats analysis, this clause is amended with a note that states that guidance on conducting a security threats analysis can be found in IEC 62443, a standard for security of industrial systems which is currently being developed. The work of the committee is documented at `http://isa99.isa.org`, where also current drafts of the standard may be found.

This means, that if a safety critical system were to be certified according to IEC 61508 Ed2, then it has to be assured that this system is also analyzed for security, and that security being handled appropriately is a pre-requisite for allowing the system to be safe in the first place.

The following section gives a short overview of available threat modeling techniques for industrial system. After that, a rational why one might decide to not use any of them is given. Following that, in Section 3 an introduction to the Hazard and Operability Study (HAZOP) is given. It is a hazard analysis method that is in wide use and highly recommended in safety standards. After that, Section 4 gives a rational on why it makes sense to adapt HAZOP for security and explains how these changes were made. Section 5 then gives an example of how such a HAZOP for security looks like.

2 Threat Modeling for ICS

The available methods for threat modeling for industrial control systems (ICS) can be divided into two groups: those methods that were actually developed for the server-/office-space and now are used for ICS as well, and those methods that were especially developed for ICS.

The problem with the first group is, that they do not work well for the special circumstances of industrial systems. The networks of industrial systems do differ greatly from the server-/office-space, most notably in the following ways:

- The dynamic of nodes is very small, that is nodes that are part of the network are not exchanged with other nodes regularly, and with the exception of special roles like maintenance laptops and the like, the topology changes either almost never or in a limited number of predictable ways (e.g. consider a robot that changes between N different tools – the tools will be (dis-)connected from the network all the time, but everything else is rather static.
- Communication is periodic, and depending on the protocol at least in part defined at configuration time. This periodicity simply stems from the need to periodically send process data like measurement values and set points for actuators.
- The nodes participating in the communication perform actions are very specific to industrial systems, in contrast to server space were databases and webservers are some of the prime functions, these functions – while available on industrial systems – are mostly only for convenience on industrial nodes and the primary function has little to nothing to do with them (if they are present and configured at all).

Those methods that are developed for threat modeling of industrial systems are mostly still under development or seem to be dead (no activity on the webpage for a considerable amount of time). An example of such a methodology is TRIKE [6].

This situation resulted in the question that lead to this work which is, whether or not it would be possible to adapt existing methods used for analyzing hazards in the safety domain so they may also be used for analyzing threats in the security domain.

3 An Introduction to HAZOP

The purpose of this chapter is to give a short introduction to hazard and operability studies (HAZOP). There is a lot of literature available that gives a more detailed introduction to HAZOP, e.g. [1] not only explains how to conduct a HAZOP but also gives a lot of rationale that stems from decades of experience by the author. Also the standards of the Ministry of Defence, UK [2, 3] shall be mentioned, although DEF STAN 00-58 has been canceled by the Ministry of Defence back in 2004 according to http://www.dstan.mod.uk, it is still a viable source for information if you are new to HAZOP. Of course this section cannot contain all the information you get from the cited documents, but should suffice to give an inexperienced reader a quick introduction to HAZOP. In addition also IEC 61882 **??** gives guidance on the usage of HAZOP.

3.1 Initiating a HAZOP Study

Assuring the safety of a product is a task that stretches all over the development life cycle of the product. In most cases, the need for it is not only a good practice to improve the quality of the product,

but a requirement by the authorities in order to assure that the users of the resulting product won't be exposed to a higher risk than necessary.

In order to being able to assess the risk imposed on the users by the system, the potential hazardss have to be identified, their risk assessed and appropriate counter measures have to be taken. As HAZOP is one of the possible methods that can be appropriate (choosing the right method is a complex task), the process of initiating a HAZOP has to be done very early in the development life cycle. The study is initiated by a *Study Initiator*, usually this is some senior manager, or a company wide safety officer. The only restriction on the person of the Study Initiator is, that it should be someone who is not directly involved in the development of the product.

One of the first tasks of the Study Initiator, is to assign a *Study Leader*, who will be responsible for the conduction of the HAZOP studies. The preparation work for study is distributed between those two.

3.1.1 Choosing a Team

As mentioned above, HAZOP is done in a team, the team is either chosen by the Study Initiator, but he could also decide to delegate this task to the Study Leader.

Study Leader The study leader is responsible to plan the HAZOP studies (at which points in the DLC, the dates, the duration, use representations, etc.) and as the name says he will lead the studies, and make sure that they are conducted in an effective and efficient way.

Designer The designers task is to explain the design to the other team members.

User / Intended User The users view is also important to identify hazards, the user or intended user is usually a future operator, or someone who has been operating similar systems in the past.

Expert The expert's task, is to help the team members to analyze the causes. A expert might be only of help for a subsystem, and therefore only present for this subsystem where he really is an expert.

Recorder The recorders job is to document the findings of the HAZOP study. This documentation has to be finished at the study, and signed by the team members. The Recorder also has to manage the documents of previous HAZOP studies, as those might be of use for follow up studies.

The choice of team members with the necessary experience and expertise is a crucial part of the HAZOP study initiation and the study initiator/leader has to make this choice very carefully. It is also very important, that all selected team members are familiar with the chosen design representations, etc. If they are not, they have to be introduced to them before the first study session, to allow them to fully participate in the study.

The different background of the members (e.g. User vs. Designer vs. Expert) is what ensures a high coverage over the potential hazards in the system. The higher the diversity in the team members background, the higher the confidence in the outcome of the study.

3.1.2 Choosing Appropriate Representations

One of the most important actions to be taken by the Study Leader **before** the studies are conducted, is to choose the appropriate design representation(s).

Furthermore, the Study Leader has to make sure, that all the team members are familiar with the chosen design representations, and if they are not give them a introduction, in order to make sure that the HAZOP study can be conducted without any unnecessary disturbances due to questions of team members not knowing the design representation.

3.1.3 Guide Words

*"**Guide Word** Defined as a word or phrase which expresses and defines a specific type of deviation from design intent." [2], section 1, 4.1.10*

A generic list of such guide words is given by [2], this list contains the following guidewords / meanings of those guidewords:

No – This is the complete negation of the design intention. No part of the intention is achieved and nothing else happens.

More – This is a quantitative increase.

Less – This is a quantitative decrease.

As well as – All the design intention is achieved together with additions.

Part of – Only some of the design intention is achieved.

Reverse – The logical opposite of the intention is achieved.

Other than – Complete substitution, where no part of the original intention is achieved but something quite different happens.

Early – Something happens earlier than expected relative to clock time.

Late – Something happens later than expected relative to clock time.

Before – Something happens before it is expected, relating to order or sequence.

After – Something happens after it is expected, relating to order or sequence.

One important point made in [1] is, that studies tend to introduce new words when this is not necessary. The authors state, that despite their decades of experience they have barely every needed any other guide words than the generic ones. According to them not the guide words itself are the essential part, but the combination of guide words and their meaning, therefore it is very important to list the guide words that are going to be used plus their meaning for this HAZOP study at the beginning of the study. This also clarifies the meaning to all the team members and prevents a different understanding of a guide word by the team members.

3.2 When to conduct HAZOP

The points in the development life cycle, when a HAZOP study should be conducted depend on the criticality of the system. For systems with a low safety-critical relevance, it might be sufficient to do it just once at the beginning of the project, for systems with high safety-critical demands, it will be necessary to do HAZOP studies at different points in the safety life cycle. For example, it could be necessary to conduct a HAZOP study to produce a preliminary list of hazards as soon as the High-Level Design is available and another one as soon as the detailed design has been done. Then the hazards identified from the detailed design can be checked against the preliminary list of hazards derived from the high-level design.

The points in time when to do the HAZOP studies is defined at the start of the project, however it might be necessary to do additional HAZOP studies e.g. when the systems environment changes.

4 Adaption of HAZOP

In Section 3.1.3 the guidewords that are usually used in HAZOP studies were introduced. Since HAZOP has its origins in safety, the definition of these generic guidewords talks about a *"deviation from design intent"* assuming this deviation to be either due to random faults or due to a systematic fault introduced during design and development of the system. However, in the following, HAZOP is used as a method to systematically detect security threats. Therefore, this assumption has to be changed to be either:

- Exploitable design or implementation flaws that permit intentional actions to be taken by an attacker, or
- Unintentional actions taken by an employee that lead to situations where either the way for an attacker is made very easy, or the result is indistinguishable to what an attacker could have done (e.g. information theft vs. information leakage, or malware that is intentionally / unintentionally introduced with an USB-drive).

So in contrast to safety, random faults are not in the picture at all but only systematic faults and their exploitation (intentional or unintentional) are considered.

Although it was also mentioned above, that the generic list of guidewords are seen to be sufficient for almost all HAZOPs, this can only be true as long as safety is considered. Using HAZOP to analyze a different system property – security for example, an adaption of the guidewords would be necessary.

The big question however is, how to get a list of guidewords that are sufficient to cover all aspects of security. The process taken to acquire such a list is part of the full workflow from a system description to actual mitigations against security vulnerabilities depicted in Figure 1. The approach taken was to first define a generic industrial system and define further develop data-flow diagrams of the data-flow between the entities in this generic industrial system.

Furthermore, the attack surface of this generic industrial system was defined and a list of threat agents as well as the point of entrance into the attack surface for each attacker were analyzed after the list of threat agents was available. Now it was possible to analyze the actions these threat agents are taking to penetrate the system – this list of actions was then used to get an exhaustive list of keywords that cover the security property of the system. This

list was then also checked against the current draft of the IEC 62443 standard [5]. The resulting list will be presented in Section4.1.

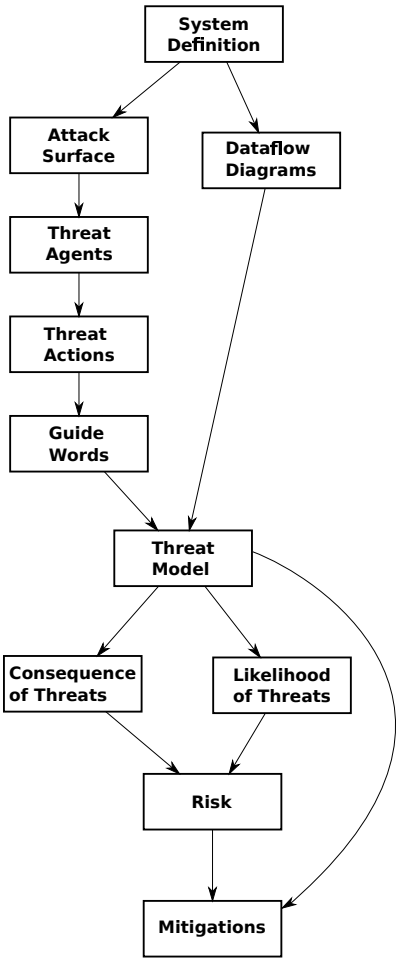

FIGURE 1: *Workflow: From a System to appropriate Mitigations*

Now the list of security guidewords can be used to perform a HAZOP study – e.g. on the data-flows in the system, represented by the data-flow diagrams. The output of this Threat Analysis is a Threat Model in the form of a list of threats as well as their causes and consequences. Ideally an indication that the threat is present or a protection against the threat are already given as well.

For all threats that do not have an detection and/or protection mechanism, the consequence as well as the likelihood shall be estimated. These two factors can then be used to calculate the risk of this threat leading to a security problem.

The risk is then used to find out for which threats additional mitigations that allow the indication of and/or protection against the threat should be found. The goal is not to eliminate all threats, it merely is to reduce the risk to an acceptable level. What level may or may not be considered acceptable is of course a question of the actual application and its environment. At last the list of necessary mitigations is available.

4.1 Security Guidewords

As already promised above, the list of guidewords that was identified during the analysis are presented now. Their use will be demonstrated in an hands-on example in Section 5.

Spoof – Is it possible for an attacker to pose as a legitimate user, or let a device under his/her control act as if it were a legitimate device?

Elevate Privileges – Is it possible for a user with insufficient permissions to perform actions he/she does not have the necessary permissions for?

Listen – Can an attacker read/listen to sensitive information?

Corrupt – Can an attacker manipulate the data in a non-systematic way (e.g. to random values the attacker does not have an influence on)? *NOTE: This can be seen as a form of Denial of Service attack.*

Tamper – Can an attacker manipulate data in a systematic way (to specific values the attacker does have an influence on)?

Denial of Service (DoS) – Can the attacker disturb communication or crash a hardware node? (Denial of Service Attack)

Malware – Is it possible for an attacker to introduce malware via this data-flow?

Wrong Connection – The attacker reconnects cables/nodes in a wrong way, e.g. so that there is a cable from an untrusted network to a trusted port, used by the application. Could this wrong connection lead to an insecure situation?

5 Example: Digital Dead Man's Switch

In the following section, a threat model for an example application will be performed, but of course this sample application has to be defined first. This is

done in Sections 5.1-5.3, the actual analysis will be done in Section 5.4.

5.1 Functional System Definition

The example system used in the following is an application called "digital deadman switch". It is a system that implements safety (see Section 5.2) as well as security (see Section 5.3 functions and thus has both: security and safety requirements.

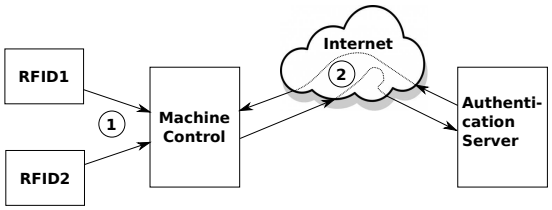

FIGURE 2: *Example Application: Digital Deadman Switch*

The idea behind the digital deadman switch is to implement a mechanical deadman switch as used in many machine applications and replace it with a digital version of that which allows further checks – not only relevant to safety but also for security. Physically, the mechanical deadman switch is replaced by two radio-frequency identification (RFID) readers which are in-ground (in the floor) of the factory. The counter part – commonly referred to as RFID tags – are integrated into the workers boots. Each worker has two RFID tags, one in the left and one in the right boot. If the worker wants to use the machine, he/she has to assure, that the left foot is on the RFID reader for the left foot (RFID_L) and that the right foot is on the RFID reader for the right foot (RFID_R).

A block diagram of the system is given in Figure 2, and the functionality of the entities in the system is as follows:

RFID_L / RFID_R These are two RFID readers that do nothing else, but periodically send the ID they are currently reading to the Machine Control.

Machine Control This is the actual controller of the machine (in this example, the actual purpose of the machine is not relevant and thus not defined). The controller only allows the operation of the machine if a valid reading from RFID_L and RFID_R was available (more on what that means in Sections 5.2 and 5.3).

Authentication Server Since it is not desirable that a company has to update permissions per machine, but at least factory wide or ideally company-wide, the Machine Control node has to check the operators credentials (that is, the two IDs of the RFID tags) against the companies database. In this example, a company with more than one factory is assumed, and the server this request has to be sent to is connected via an open network – the Internet.

Note, that the authentication server only is used for authentication of users, not devices.

The above description gives a very high-level overview of the system, on this high level only two data-flows are considered labeled as ① and ②. These labels will be used in Annex A, to refer to those to data-flows. Next the properties for safety and security that shall be reached by this system are considered.

5.2 Safety Functions

From a safety perspective the replacement of the mechanical deadman switch by this digital deadman switch is rationalized with the following advantages:

- Tricking the safety function becomes a lot more complicated. The mechanical switch could be tricked by simply putting a stone on it.
- In addition to detection the presence of the operator the direction into which the operator is looking can now be detected (very course grain, but still) by considering the fact that if the operator is looking into the wrong direction (away from the machine), then then the left foot will be in the position where the right foot should be and the other way around. This can be detected by the machine control and the machine can be kept in the standby mode instead of being activated.
- By looking up if the operator is allowed to use that particular machine, it can also be checked if the operator actually has the training necessary to use the machine. If the necessary training has not been given to that operator he/she is not allowed to use that particular machine.
- In addition to the check of the operators training, the working hours may be observed, in order to assure that the operator does not work for too many hours without taking a break.

- In some countries, the wearing of protective gear is not enforced rigorously, using this kind of system where the RFID tags are integral parts of the shoes will enforce this (at least to some extend).

5.3 Security Functions

In addition to the safety functions described in Section 5.2, the digital deadman switch also provides security functions:

- Who is the operator?
- It is assured, that no unauthorized person can operate the machine.
- When checking the working hours of the operator, it can also be assured that the operator is not (mis-)using the machine for private business outside of his/her shift.

What is interessting here is, that the first two of these security functions overlap with the safety functions – this is not very surprising as safety and security are interdependent system properties.

5.4 Threat Model

Now, as the system under consideration was described in the previous subsections, it is time to establish a threat model using the HAZOP analysis method adapted for security with the list of guidewords as listed in Section 4.1.

The analysis is done in tabular for, since this table is rather big it is moved to Annex A at the end of this paper. In the following it won't be necessary to go through all the HAZOP items of the example, but some representative examples shall be discussed in short:

HAZOP Item 1: This item represents the best case: although the keyword unveils a threat, that could bring the system into trouble, but there is already a protection in place that takes care of this threat.

HAZOP Item 2: Sometimes (depending on the context) different keywords can have the same meaning. In those cases only a reference to the relevant HAZOP item is given. In this case, the keyword *Elevate Privileges* is interpreted as an attacker going from not being part of the conversation to being part of the conversation.

HAZOP Item 7: Depending on the context it may also be possible that a keyword is not applicable at all. In the example, the communication is bound to a very specific message that is sent periodically. This can not be misused to upload malware to the machine control node.

HAZOP Item 12: This item actually uncovered a problem for which no mechanism was present before. For this item it is now necessary to either:

- Argue why no protection or even detection mechanism is necessary. In example, it could be argued that it is incredibly improbable that replacing the encrypted message with random data will result in the unencrypted message being meaningful (e.g. a valid pair of RFID IDs).
- Introduce a new mechanism to detect that the message was corrupted – e.g. add a checksum that is calculated before encryption to the message.

This selection shows us how this systematic analysis of system will reveal all the present threats. For some of these a mitigation will already be presents, others won't be relevant at all and for others it will be necessary to introduce a new mitigation.

Of course the above given example is a very high-level example, and it will be necessary to go into detail and analyze more elaborate data-flows as well. This is one of the core concepts – do a first analysis soon in the development process on a very high level, do more analysis as you go into detail on your system design, until at the end you analyze the implementation itself.

6 Conclusion

The work presented in this paper shows, that it is possible to re-use the well known methods from the safety domain and – with slight adaptions – re-use them for the security domain.

The huge advantage of this being, that only one method has to be known within the development team. At least for some members of the HAZOP team this means that only one instead of two methods need to be known. Looking at the members of a HAZOP team as described in Section 3.1.1, in example the *Study Leader* , *Recorder* and *(Intended) User* will have the exact same role in the HAZOP with the only difference being the list of guidewords that is used.

Other roles, like the *Expert* will have to be filled with different background – a safety engineer if hazards shall be analyzed, a security engineer if threats shall be analyzed.

The fact that HAZOP is a well known methodology with a long track-record. The adaption as presented above are important for making it possible to even analyze security properties. The way of how those guidewords for security were determined is well documented (with a short summary of how it was done in Section 4).

This allows a thorough argumentation of why this method is chosen for threat modeling and why it is seen as suitable. This result is not very surprising, as both – safety and security – are system properties. Although they are often treated separately, the interdependencies between them make it reasonable to use the same methods, for economic as well as technical reasons. Above, it was shown that this approach is not only reasonable, but that it is also possible to use a hazard analysis method for threat modelling.

References

[1] *System Safety: HAZOP and Software HAZOP*, Felix Redmill, Morris Chudleigh and James Catmur, 1999, Wiley

[2] *Defence Standard 00-58 - HAZOP Studies on Systems Containing Programmable Electronics, Part1: Requirements, Issue 2*, Ministry of Defence, UK, 2000

[3] *Defence Standard 00-58 - HAZOP Studies on Systems Containing Programmable Electronics, Part2: General Application Guidance, Issue 2*, Ministry of Defence, UK, 2000

[4] *IEC 61882:2001 Hazard and operability studies (HAZOP studies) - Application guide*, 2001

[5] *IEC 62443 (DRAFT): Security for Industrial Automation and Control Systems*, isa99.isa.org, 2015

[6] *Webpage of the TRIKE threat modeling methotology*, http://octotrike.org/, 2012

A HAZOP Table for the Example

HAZOP Item	Item	Guide Word	Cause	Consequence/ Implication	Indication/ Protection	Question/ Recommendation
1	①	Spoof	The attacker replaces the RFID1/2 by his own device(s).	The attacker is able to send any combination of RFID1 + RFID2 to the machine control unit.	The RFID1/2 node and the machine control node authenticate each other at boot-up and exchange a session key, which is used to encrypt the exchanged data.	Note: the exchanged data includes current timestamps which results in a variation of the messages.
2	①	Elevate Privileges	Same as HAZOP item 1.			
3	①	Listen	The attacker listens to the communication between RFID1/2 and the machine control and reads the valid pairs of IDs from RFID1 and RFID2.	The attacker collects valid pairs of RFID IDs that can be used for a later attack on this machine or other machines in the factory.	As stated in HAZOP item 1, the communication between RFID1/2 and the machine control node is encrypted.	

4	①	Corrupt	The attacker corrupts communication between RFID1/2 and the machine control.	The RFID IDs read by RFID1/2 can not be communicated to the machine control any more.	The machine control expects a periodic message with an ID (or no ID if no tag is present). If theses periodic messages from RFID1/2 are not received the machine control knows that something is wrong.	Basically this is a DoS attack.
5	①	Tamper	The attacker performs a man-in-the-middle attack and alters the IDs received by RFID1/2 before passing the message on to the machine control node.	The attacker is able to make a) legitimate IDs into random non-legitimate numbers b) make legitimate IDs into the "no RFID tag present" message c) make an non-legitimate ID into an legitimate ID	a) If IDs that are not in the database are detected are received, this is an indication that something bad might be going on and an alert is triggered. b) and c) Encryption of messages (see HAZOP item 1 for details).	
6	①	Denial of Service (DoS)	Same as HAZOP item 4.			
7	①	Malware	Not applicable.			Only well specified messages are accepted.
8	①	Wrong Connection	The attacker exchanges the cables of RFID1/2.	The IDs for Left and Right foot will always be exchanged.	This will result in always having ID pairs not in the database which will be detected and a security alert will be raised.	
9	②	Spoof	The attacker poses as the authentication server.	For each pair of IDs the attacker can tell the machine control that it is a valid pair of IDs.	The motor control and the authentication server authenticate each other with pre-configured keys.	
10	②	Elevate Privileges	Same as HAZOP item 9.			

11	②	Listen	The attacker listens to the communication between machine control and the authentication server.	The attacker collects valid ID pairs by listening to the communication.	The communication is via an open network, strong encryption is used to prevent that someone listens to the communication.	
12	②	Corrupt	The attacker manages to corrupt messages between motor control and authentication server.	The corrupted messages may or may not have a meaning to the receiver. If it does, a wrong behaviour may be the result (valid IDs instead of invalid).	No mechanisms to detect such an event.	
13	②	Tamper	The attacker replaces the messages content.	The attacker is able to control the content of the messages that are passed between the motor control and the authentication server.	This is very unlikely as the messages are encrypted and it would require the attacker to break the encryption first.	Note: A state of the art crypto-algorithm and key-length shall be chosen.
14	②	Denial of Service (DoS)	The attacker prevents a connection between the motor control and the authentication server.	No exchange of messages between motor control and authentication server is possible anymore.	The loss of connection between the machine control and the authentication server results in a security alarm.	
15	②	Malware	Not applicable.			There is a well defined protocol in use – no upload of data possible.
16	②	Wrong Connection	The attacker disconnects the machine control (or the authentication server) from the Internet.	The machine control can no longer check ID pairs for their validity at the authentication server.	This is detected by the machine control and a security alarm is raised.	