Cyber Security of the Railway wireless system: detection, decision and Human-in-the-Loop

Christophe Gransart, Virginie Deniau, Eric Pierre Simon, Anthony Fleury, Stéphane Lecoeuche, Patrick Millot, Emilie Masson

▶ To cite this version:

Christophe Gransart, Virginie Deniau, Eric Pierre Simon, Anthony Fleury, Stéphane Lecoeuche, et al.. Cyber Security of the Railway wireless system: detection, decision and Human-in-the-Loop. TRA 2018, 7th Transport Research Arena, Apr 2018, Vienne, Austria. 10p. hal-01852410

HAL Id: hal-01852410
https://hal.archives-ouvertes.fr/hal-01852410

Submitted on 1 Aug 2018

Proceedings of 7th Transport Research Arena TRA 2018, April 16-19, 2018, Vienna, Austria

Cyber Security of the Railway wireless system: detection, decision and Human-in-the-Loop

Christophe Gransart[1 5], Virginie Deniau[1 5], Éric Pierre Simon[2 5], Anthony Fleury[3 5], Stéphane Lecoeuche[3 5], Patrick Millot[4 5], Émilie Masson[5*]

[1]Univ Lille Nord de France, IFSTTAR, COSYS, LEOST, F-59650 Villeneuve d'Ascq, France
[2]University of Lille, IEMN lab
[3]IMT Lille Douai, Univ. Lille, Unité de Recherche Informatique Automatique, F-59000 Lille, France
[4]University of Valenciennes and Hainaut Cambrésis, LAMIH CNRS UMR 8201
[5]Institut de Recherche Technologique Railenium, F-59300, Famars, France

Abstract

The networks used in the Railway domain are usually heterogeneous, not enough protected and not fitted to the usual Cyber Security requirements in terms of sustainability, protection and attack detection. Furthermore, the quick evolution of the telecommunication means, the threats and the sustainability aspects have to be taken into account in order to protect the Railway system.
The paper presents the first contributions on Cyber Security for railways that can be divided into three main aspects dealing with the Cyber Security of the wireless part of the railway communication system: detection, decision and Human-in-the-Loop. Part of the work will be devoted to the development of an Open Pluggable Framework (OPF). The OPF is a software framework based on automation principles. It monitors the environment, then some algorithms detect abnormal behaviours, and next, OPF decides which reaction to take and finally apply this action (e.g. an alarm or a reconfiguration). The last part "human in the loop" aims at answering the questions: what happens if the automatic countermeasures fail and how the driver can cope with the attack consequences. It consists in placing professional drivers and Central Traffic Control operators in a realistic simulator and playing scenarios involving attacks and observing the reactions of the professional drivers. A preliminary methodology is proposed and discussed through a concrete case study.

Keywords: cyber-attacks detection, Software Defined Radio (SDR), classification algorithms, Train-tram Simulator, scenarios of threats, analysis of driver behaviour

1. Introduction

The networks used in the Railway domain are usually heterogeneous, not enough protected and not fitted to the usual Cyber Security requirements in terms of sustainability, protection and attack detection. Furthermore, the quick evolution of the telecommunication means, the threats and the sustainability aspects have to be taken into account in order to protect the Railway system.

Through the SmartRaCon consortium, Railenium is an Associate Member on Innovation Programme (IP) 2 of the Joint Undertaking Shift2Rail. The activities on IP2 started through the X2Rail-1 project that deals with the start-up activities for Advanced Signalling and Automation System. In particular, Work Package (WP) 8 deals with the Cyber Security issues.

The paper presents the first contributions of Railenium on WP8 that can be divided into three main aspects dealing with the Cyber Security of the wireless part of the railway communication system: detection, decision and Human-in-the-Loop. Part of the work will be devoted to the development of an Open Pluggable Framework (OPF). The OPF is a software framework based on automation principles. It monitors the environment, then some algorithms detect abnormal behaviours, and next, OPF decides which reaction to take and finally apply this action (e.g. an alarm or a reconfiguration).

The first part of the work deals with the development of generic detection solutions able to detect different types of cyber-attacks targeting different levels of the network stack. The detection algorithms will be implemented on

[*] Corresponding author. Tel.: 33 (0)6 83 11 34 57
E-mail address: emilie.masson@railenium.eu

Software Defined Radio (SDR) cards in order to propose hardware probes monitoring the physical layer. Software probes solutions will then be developed to monitor activities on the other levels of the network stack for different network protocols. The output of these probes will be injected into the OPF which constitutes the monitoring architecture.

The second part of the work consists in investigating the use of adaptive classification algorithms for the detection and identification of attacks. It would allow to: (1) Detect unknown (new) internal and external threats and intrusions, (2) Build models with incomplete knowledge about the normal and safe modes, (3) Adapt the built models to evolving behaviours of the attackers to break the security rules. In such a system, some basic rules are coded in the first initialisation of the system and then the attack detection system will monitor and analyse any kind of possible drift in the behaviour of the operator to detect and localize further attacks more precisely.

Finally, Human factor has to be addressed in the risk analysis. It consists in assessing the professional human driver and CTC (Central Traffic Control) supervisor abilities to react to (simulated) cyber-attacks, or to their consequences, in a realistic simulator, by reproducing scenarios involving humans, analysing their behaviours and their abilities to detect and to mitigate the threats. Then, the method will look for strategies, indicators and devices useful for human counter measures. More generally, the work consists in improving human detection and recovery and enhancing their system resilience.

2. Detection solutions of cyber-attacks

Given the rapid evolution of telecommunication and cyber threats, the railway sector has a double concern to evolve to improve its services and to protect itself in order to continue to guarantee its safety.

As part of the X2Rail-1 project, Cyber Security is being considered through the design of an Open Pluggable Framework (OPF), which is responsible for managing surveillance data, decision-making functions and activation of adapted countermeasures.

Work in progress is based on accurate analyses of the impacts of attacks on different communication standards. This work aims at identifying the most relevant parameters to be monitored for detecting and classifying the attacks. Different jamming attacks, with different waveforms, have to be processed in order to be able to detect the greatest number and to differentiate their respective impacts.

Our approach is presented in more details in the following. We started our analysis with jammers bought on the Internet. These jammers constitute a true threat since they are cheap and very easily accessible to the general public.

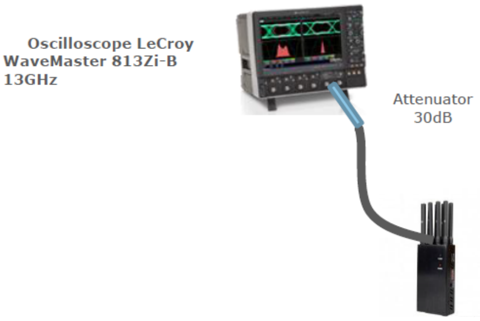


Fig. 1 Characterization of the signals produced by jammers accessible to general public

The first step of our analysis was the characterization of the jamming signals. This has been carried out by recording a sequence of the signal with an oscilloscope (see Fig. 1). Then, a time-frequency transform tool has been applied (see Fig. 2).

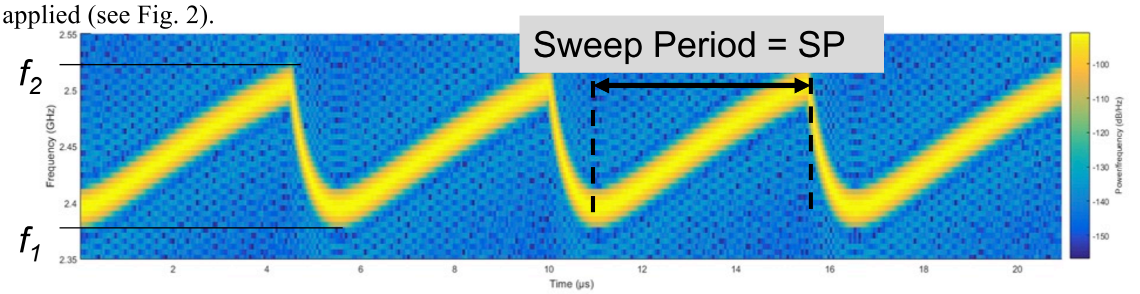


Fig. 2 Time-frequency representation of a jamming signal

It can be observed that the signal is as frequency sweeping signal, with a period (denoted in the following by the acronym SP for sweep period) of 5.5 us.
Several jammers accessible to general public have been analysed and it turned out that it was always the same kind of signal. Furthermore, this result is confirmed by the outcomes of the FP7 SECRET project, where several jammers were characterized. The only difference between the jamming signals was the value of the SP, which varies from 5.5 us to 8 us behind the jammers that we tested.
Then, the impact of the jamming attack on the performance of a Wi-Fi transmission 0 is investigated through measurements carried out in an anechoic chamber 0.
As mentioned earlier, the nature of the jamming signal keeps the same, but the value of the SP can vary. We have then generated jamming signal in making varying their sweep period in order to investigate in detail the role of this SP parameter.

2.1. EM attack on a WIFI communication applying a frequency sweeping interference signal

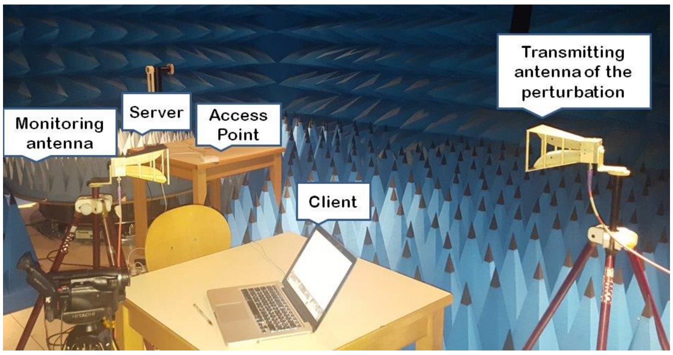


Fig. 3 Experimentation with a 802.11n communication in presence of jamming signal

In order to study the jamming signal impact on the IEEE 802.11n communication network performance, we selected an interference signal with a defined SP and we varied the Interference/Signal Ratio (ISR) by progressively increasing the interference signal power.

$$ISR = P_I/P_S$$

where P_I is the interference jamming signal power and P_S is the Wi-Fi signal power. P_I and P_S are obtained from oscilloscope measurements.
We present, in Fig. 4, the minimum level of ISR required to loose the communication, as a function of the SP.

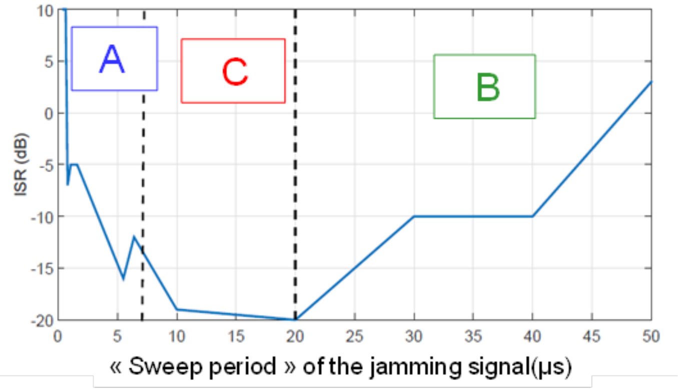
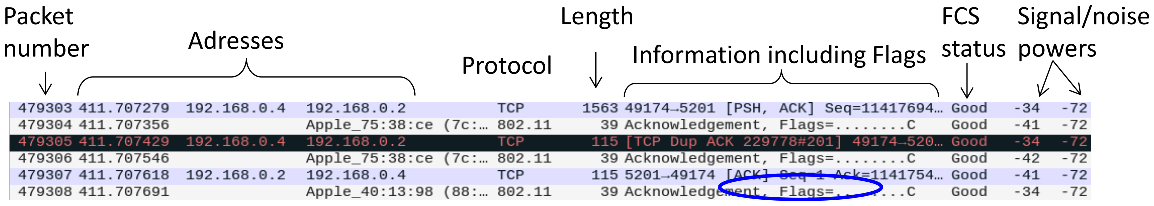


Fig. 4 Required value of the ISR to interrupt the communication, as a function of the SP

We noticed three different behaviours in Fig. 4, represented by the Area A, B and C in the figure. The modulation of the Wi-Fi, the OFDM modulation (orthogonal frequency division multiplexing), plays a crucial role for the interpretation of these results, as is presented below.
A: For a low SP, the FFT of the OFDM receiving stage modifies the distribution of the interference signal in the frequency domain. At the OFDM receiver stage, the interference signal appears as a series of frequency harmonics spaced by 1/SP. So, only some subcarriers are affected, then inducing isolated errors.
B: Since the SP is higher, the interference only covers part of the 20 MHz channel over the duration of an OFDM symbol.
C: All subcarriers are affected and interference covers the entire band of the channel. It corresponds to the worst case.
Finally, we noticed that according to the sweep time, the interference signal power requires to corrupt the communication can vary of 30 dB.
The analysis of commercial jamming was the first step. However, other threats come from the proliferation of EM signal generation equipment such as Software Defined Radio, communication shields and modules and antennas for instance. Now any jamming signal can be created with few efforts, and future standards will have to face this new threat.

2.2. Detection of attacks

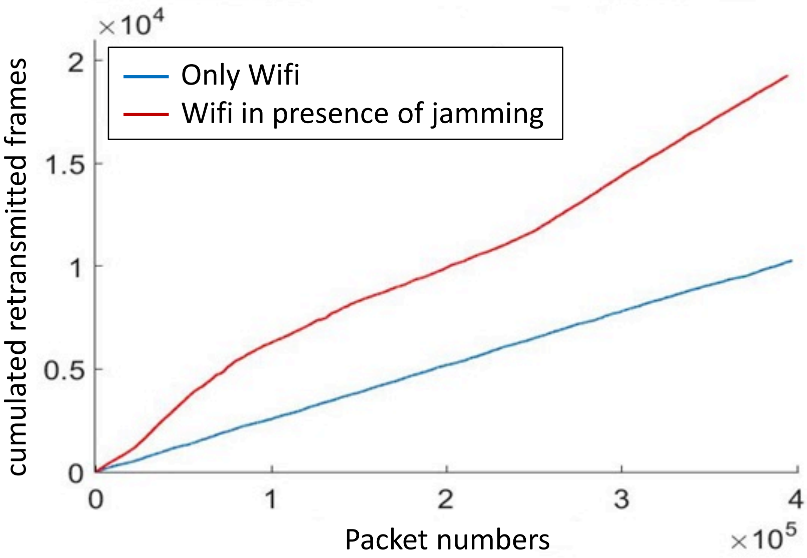
To countermeasure a cyber-attack situation, the first step is to be able to detect such attack. Different detection solutions based on the monitoring of the physical link were studied in the project SECRET 0. In this TRA communication, we illustrate the monitoring approaches which can be developed on the data link layer.
The monitoring approach is based on the use of the Wireshark Software. Wireshark is a network protocol analyser. It allows us to record all the traffic exchanged by the different computers using the wireless link. We obtain the frames from the Data Link layer (OSI layer 2) up to the Application layer (OSI layer 7). Currently, we are monitoring signalling frames from the IEEE 802.11n protocol and also segments from the TCP layer.



```
479303 411.707279  192.168.0.4  192.168.0.2           TCP     1563 49174→5201 [PSH, ACK] Seq=11417694… Good  -34  -72
479304 411.707356               Apple_75:38:ce (7c:… 802.11    39 Acknowledgement, Flags=........C   Good  -41  -72
479305 411.707429  192.168.0.4  192.168.0.2           TCP      115 [TCP Dup ACK 229778#201] 49174→520… Good  -34  -72
479306 411.707546               Apple_75:38:ce (7c:… 802.11    39 Acknowledgement, Flags=........C   Good  -42  -72
479307 411.707618  192.168.0.2  192.168.0.4           TCP      115 5201→49174 [ACK] Seq=1 Ack=1141754… Good  -41  -72
479308 411.707691               Apple_40:13:98 (88:… 802.11    39 Acknowledgement, Flags=........C   Good  -34  -72
```

Fig. 5 Illustration of Wireshark

By defining specific filters, we extract some information from Wireshark and we convert this information in numerical attack indicators. To illustrate the approach, the Fig. 6 represents the cumulative curve of the retransmitted frames in case of presence of jamming signal and without jamming signal.



We obtained a separation of the space which can permit to envisage classification works based on these retransmitted frames indicator. Other indicators are currently under study. We are notably studying an indicator based on the Frame Check Sequence (FCS) which permits to indicate if the received frame is corrupted.

3. Attack classification

From the data that have been acquired previously, some "simple" features have been extracted to check whether it is possible to find some reproducible threshold separating the space between the case of a jamming and of a normal use. It is the case of a first analysis on the number of retransmitted data as presented in the Fig. 7.
However, our next goal is to find more generic models that could be able to differentiate two classes, an attack (whatever the attack) and a normal use of the network. For that, we will consider the same data as previously presented but we will extract more features from them. The general architecture of the proposed system is presented on Fig. 7.

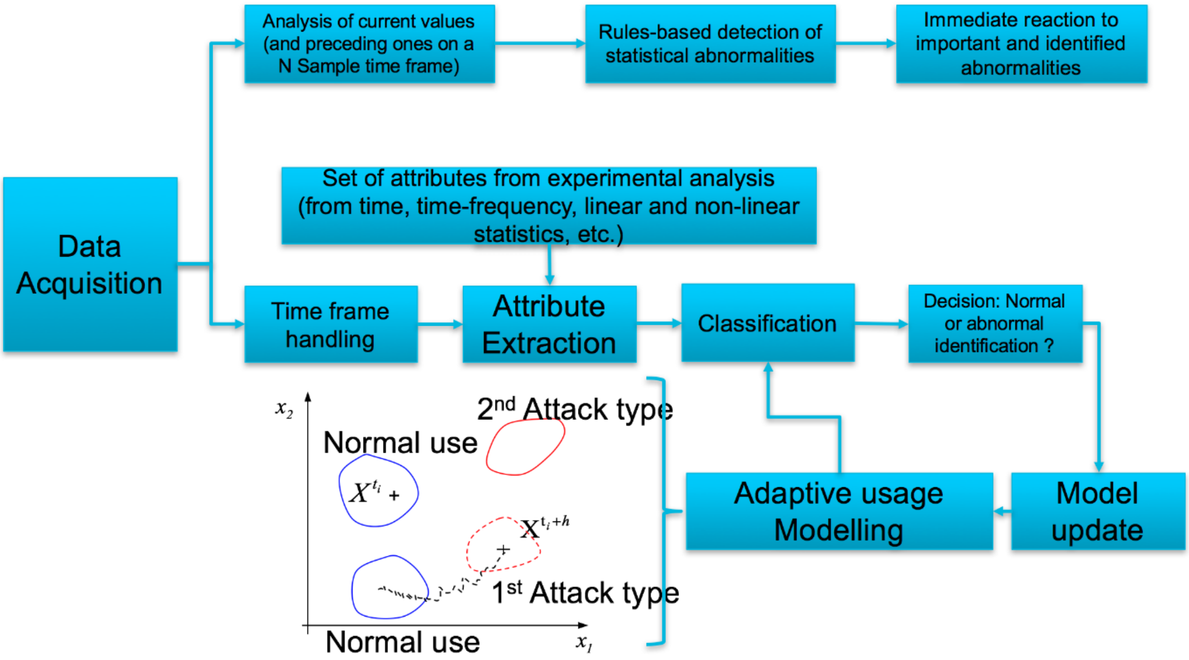

Analysis of current values (and preceding ones on a N Sample time frame)
Rules-based detection of statistical abnormalities
Immediate reaction to important and identified abnormalities
Set of attributes from experimental analysis (from time, time-frequency, linear and non-linear statistics, etc.)
Data Acquisition
Time frame handling
Attribute Extraction
Classification
Decision: Normal or abnormal identification ?
Adaptive usage Modelling
Model update
x_2
2nd Attack type
Normal use
X^{t_i} +
X^{t_i+h}
+
1st Attack type
Normal use
x_1

Fig. 7 Proposed architecture for attack identification from different signals and data.

This architecture contains two main parts. A first part considers the data using a time window approach and will compute some very simple features that can easily detect some specific misuse. One of these feature is the previously presented one that will detect jamming. This case is not the worth but for some of the attacks, a real danger could appear. For these, we have to react as soon and as easily as possible (by for instance closing a connection or turning off a controller). For them we need to compute some very simple features and that do not need a large delay in the data to be evaluated correctly.
The second kind of danger that will be handled are less vital attacks that could occur. For them, we propose to compute some features on the data and, using an adaptive classification method, to detect normal and attack modes in the signal. For the features that will be extracted, we propose to evaluate the capabilities of the following ones:

- Statistics on the data given by Wireshark:
    - Number of packets that are emitted from one source to another in a given time and their length;
    - FCS status of each transmissions;
    - Extraction from the flags;
    - Variability of the SNR on a duration.
- Time-Frequency analysis on the raw signals: some of our past and present works focused on analysing time-frequency content of signals including the detection of the number of modes that they contain and their location in frequency, transitory effects, stationarity, regularity of predictability. This has been

done using time-frequency representation based on wavelets or on Fourier transforms, and using indexes such as information entropy and their adaptation to non-stationary signals. Such indexes can easily be adapted to the signals that we acquired to detect short or long-term changes in their behaviour.

From these features, we then plan to apply machine learning techniques and specifically adaptive classification algorithms. We worked on algorithms that are able to detect some changes in the data and create new classes on-the-fly if necessary. For the beginning of the process, we will identify some attacks that we were able to reproduce in experimental conditions such as what have been described before. Then, the algorithm will analyse the data as they arrive and try to classify it as one of the classes that exist: normal or one of the known attack. If none of them is consistent with the current data, a new class will be created that will be considered by defaults as attack and will need an intervention to check.
From the data acquired, experimental results are being already processed. More data will be needed especially on the "normal behaviour" class, to perfect it and have some more interesting results.

4. Human approach for Risk Analysis and Threat mitigation

This third part aims at answering the question: what happens when the automatic countermeasures fail?

4.1. Principles

Our track of answer uses risk management methods dealing with prevention, decision-making, action taking, crisis management and recovery, taking into account consequences of unexpected events, whether the origin be a technical failure, an unwilling human error or a deliberated attack [5].

The approach globally developed in our project consists in three complementary steps: ***prevention***, where any unexpected event could be blocked or managed before its propagation; ***recovery***, when the event is close to result in an accident, making protective measures mandatory to avoid the occurrence of the accident; and possibly after the accident occurs, ***management of consequences*** is required to minimize damages, at least the most severe ones. The two first parts of this paper deal with automatic approaches for prevention and recovery. If these automatic means fail, the humans involved in the control and/or in the supervision of the transport system remain the last "barrier" to mitigate the threat. The question is then: will they be able to do that, and which assistance tools could help them to succeed in these tasks?

Our approach is focusing on ***"human(s) in the loop*****"** systems and simulations, taking advantage of human ability to cope with unexpected dangerous events on one hand, and attempting to recover from human errors and system failures on the other hand. Our competences are developed both in ***Human-Computer Interaction*** *and* ***Human-Machine System.*** Interactivity and human-centered automation are our main focuses*,* [6].

The principle consists in 1) placing the professional human drivers and PCC supervisors in a realistic simulator, face with scenarios which would result of the consequences of a simulated cyber-attack and 2) to observe them in order to assess their abilities to react.

4.2. Simulator PSCHITT-Rail

The Simulator PSCHITT-Rail is developed at LAMIH, lab of the University of Valenciennes. It allows 4 functionalities:

- *an editor of infrastructures*: railways for tram and train, with their environment (traffic, passengers on the platform, technical equipment); 45 km of railways are stored in the data base and original tracks can be developed by the software designer,

- *an editor of scenarios* allowing to create time tables, itinerary management and events. Several scenarios corresponding to unexpected events can be stored in the data bases and played several times to several human drivers or supervisors. For instance, events occurring on the *trains* could result on braking defaults, wrong on-board signals… On the *railway*, error on rail switching control device, disturbance on the way signals, disturbances on rail-road crossing barriers can be simulated. Specific events can also be programmed by the software designer. For obvious reasons of confidentiality, we do not go further in describing the simulated attacks.

- *a driving position* with one driver equipped with the same devices as in a real train or tram (see Fig 8). In front of the cabin, several screens display the dynamic scenes the simulated tram is inserted in.

- *a supervision work place* (Central Traffic Control Position) with a human supervisor face to the same devices as in a real system: network synopsis, trains signals, rail switching …

The simulator is mobile and its movements are programmed to give the cabin some realistic dynamics regarding the simulated scenario (see Fig 9).

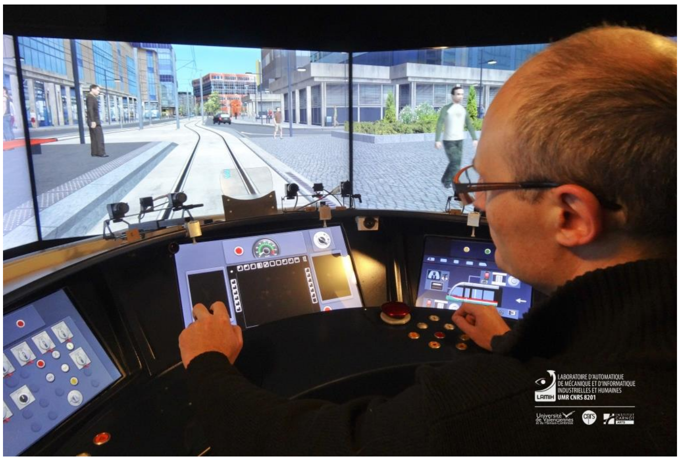


Fig 8 Simulator PSCHITT-Rail Driver position

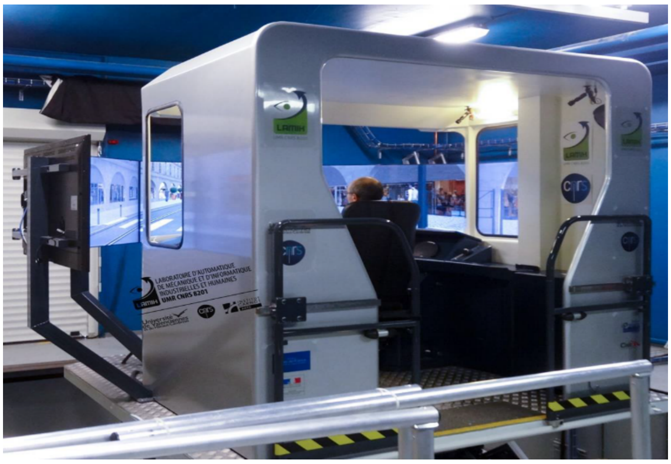


Fig. 9 Simulator PSCHITT-Rail cabin

4.3. Methodology

The study consists of different steps:

It starts *to elaborate a list of threats* and building related scenarios (ie. an external entity takes the control of the train, or of certain devices or the control of signals). The choice of the threats is done through an analysis of the vulnerability to cyber-attacks especially of connected and wireless devices. *The consequent behaviour of the related devices* is then used as a scenario to be simulated. Devices robust to cyber-attacks for instance because of their old technology are ruled out of the list.

- Then *it makes simulations of scenarios involving real humans* (usually professional drivers of a tram company) and analysing the human behaviours through dedicated sensors like camera, eye movements, completed with dedicated indicators such as workload, Situation Awareness, [6] and questionnaires, in order to *evaluate their abilities to detect the threats, and to give relevant answers*. The scenario can be replayed of line in front of the tested human operator in order to allow the analyst asking relevant questions for further explanations on the operator behaviour (called auto-confrontation). At least 10 real human operators must participate to the experiments, in order to allow the analysts comparing the different answers and detecting convergences as well as divergences in their behaviour.

- The third step consists in *analysing these answers*, the most relevant ones could be derived to establish new strategies, new procedures and/or new devices useful for human countermeasures.

- Finally, in an ideal project, these *procedures and/or devices should be integrated to the simulator* for evaluating their utility with new scenarios.

4.4 Case studies and first results

Following the methodology edicted above, the first choice was to identify *a set of vulnerable devices* the unusual behaviours of which could result of a cyber-attack, ie : 1) loss of the speed display (or display of a wrong speed value) in the cabin, 2) loss of braking, 3) jamming or disappearance of the camera rear view) in order to study the driver answers. The list of threats to be reproduced in the simulator is established according to 4 criteria : the *technical plausibility of the threat* (is it technically possible?) ; *the dangerousness of the threat* and because we are in an exploratory study, the *technical possibility to simulate* the attack with our simulator. In a more systematic approach, this last criterion can be waived in enhancing the simulator. A fourth criterion is the *possibility to transpose the attack to another kind of rail system*, for instance from tramway to metro or to train.

As an example we describe here the result of the loss of control of the rear vision which usually allows the driver to monitor through cameras the behaviour of the passengers on the platform, especially when the tramway arrives in a station. At this moment he/she selects the side corresponding to the platform for opening the doors. That allows displaying on the screens the content of what take the front and rear cameras. Then the driver can monitor the passengers entering and going out the tram.

The scenario starts with a normal driving during 30 minutes in a realistic and moderately loaded environment (down town, station with several pedestrians, rail-road crossing). Then the attack appears and results in the loss of rear vision when the tram arrives at an overcrowded station with the platform on the right side of the way. The driver must then select the right side doors to be open, but due to the attack, the rear view shows an empty platform, despite a lot of passengers were visible outside on the platform and even on the pedestrian sidewalk (through the simulator screen, see fig. 10). The usual procedure to follow by the driver in that case, is first to make the situation safe and then to call the Central Traffic Control operator for asking what to do (generally disembarking all passengers of the tram and returning to the garage for maintenance).

Then we would have logically expected that the different tested drivers apply this procedure, thinking to a breakdown. But none of the 6 drivers detected the display discordance (no extension of the stop time in the station, no particular comment, to him/herself or to the CTC operator, no special surprise). During the debriefing after the experiment all drivers say they did not detect an abnormal display of the rear views. They were all focused on the driving post showing the doors, they wait they close and then restarted the tram to quit the station.

This preliminary result is very instructive and surprising:

- many drivers are not trained for detecting such abnormal situations, they seem to favor the direct vision instead of the devices, even these ones increase the safety;
- no training sessions are proposed to the drivers to detect as well as to manage such situations, and moreover no dedicated procedures do exist.
- therefore, even the automatic countermeasures fail to prevent or to block the attacks, at least they could provide a signal in order to alert the human drivers or supervisor of the occurrence of the attack and place them in an increasing vigilance state for allowing them to react.

Of course these preliminary results must be confirmed by the other cases as evoked above, but they are encouraging for developing a relevant methodology to reinforce the driver ability to detect and treat the cyberattacks, especially to define and implement relevant devices for that, that could be seen as “manual countermeasures”. Moreover this methodology could be extended to several other situations for reinforcing trouble shooting procedures for the rail driving and management.

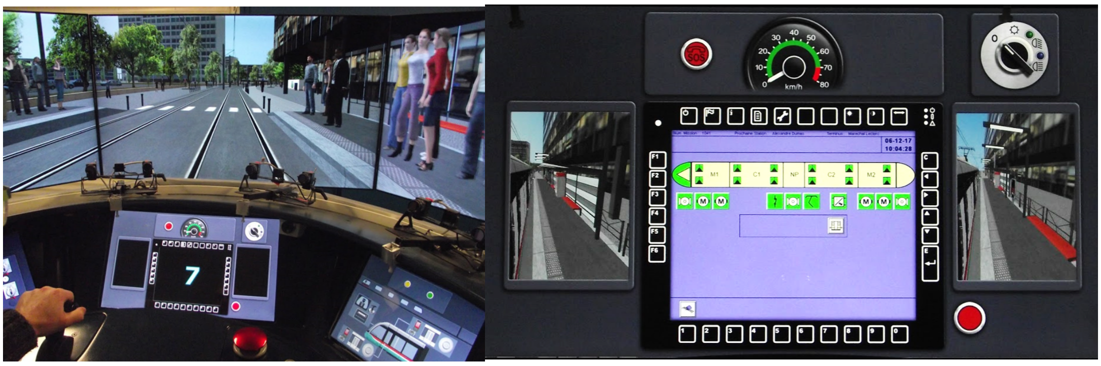


Fig. 10: Example of rear view display (right part) incoherent with the real direct outside view (left part)

5. Conclusion

The first part of the paper presented our approach to investigate the impact of commercial jammers on a Wi-Fi transmission.
First, the nature of the jamming signal has been identified through measurements and time-frequency analysis. The jamming signal has been modelled as a sweeping-frequency signal.
Second, the jamming signal has been generated with an arbitrary signal generator, and measurements have been performed in an anechoic chamber to assess the impact of such signal on the quality of the Wi-Fi transmission.
The goal of this study was to identify the most harmful interference factors on the Wi-Fi communication, with then the goal of paving the way for the design of a receiver able to adapt itself to become resilient to this kind of interference.
It turned out that the relationship between the sweep period and the time-window duration of receiver signal process played a crucial role. An interesting perspective of this study would be to investigate other factors from higher network layers.

When an attack is detected it is important to deploy countermeasures as early as possible. To accelerate the deployment speed of countermeasures, the work in progress is aimed at deploying its automatically towards jamming attack. We plan to deploy an adaptive classification algorithm based on machine learning techniques to give automatic preventive countermeasures to normal or one of the known attacks and create an alert class which require an audit in case of changes in the data.

The human approach is developed to give a final answer to consequences of cyber-attacks especially if the preventive and recovery automatic countermeasures fail. In these cases, the human driver or supervisor is alone to face with the attack and we want to study experimentally the strategies he/she develops to answer to the attack. A methodology has been described involving the way to choose the more realistic attacks, to simulate their consequences in dedicated scenarios played on a realistic dynamic simulator and the data to be recorded during the experiments and the expected results. An example shows the experimental results of the loss of the rear view when the tram arrives at a station and when the passengers enter and go out the train. The behaviour of 6 drivers were analysed and showed they did not detect the malfunction. We briefly discussed these first results and propose to extend the methodology to more general trouble shouting situations in order to define new operation procedures, especially in the cases they do not exist at present.

Acknowledgements

This work was performed in the framework of the WP8 (Cyber Security) of the X2Rail-1 project. The authors would like to thank the Shift2Rail Joint Undertaking for financial support. The authors would also like to thank the engineers of IRT Railenium, rHesuS consulting and LAMIH-PSCHITT-Rail for their involvement.

6. References

[1] “IEEE standard for information technology–telecommunications and information exchange between systems local and

metropolitan area networks–specific requirements part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications - redline," IEEE Std 802.11-2012 (Revision of IEEE Std 802.11-2007) - Redline, pp. 1– 5229, March 2012.

[2] V. Deniau; C. Gransart; Grecia L. Romero; E. P. Simon; J. Farah, IEEE 802.11n Communications in the Presence of Frequency-Sweeping Interference Signals, in IEEE Transactions on Electromagnetic Compatibility, Year: 2017, Volume: PP, Issue: 99, Pages: 1 - 9, DOI: 10.1109/TEMC.2017.2684428,IEEE Early Access Articles.

[3] S. Mili, V. Deniau, D. Sodoyer, M. Heddebaut, S. Ambellouis, Jamming Detection Methods to Protect Railway Radio Communication. International Journal of Engineering and Innovative Technology (IJEIT) Volume 4, Issue7, January 2015, pp 71-77.

[4] H. A. Boubacar, S. Lecoeuche, S. MAouche. SAKM: Self-adaptive kernel machine A kernel-based algorithm for online clustering. Neural Networks, 2008, vol. 21, no 9, p. 1287-1301.

[5] P. Millot, (ed). Risk management in Life Critical Systems, ISTE-Wiley, London. October 2014, 420 pages, ISBN: 978-1-84821-480-4.

[6] P. Millot, (ed). Designing Human-Machine cooperation systems, ISTE-Wiley, London, June 2014, 386 pages, ISBN 978-1-84821-685-3.