Security for a Multi-Agent Cyber-Physical Conveyor System using Machine Learning

Gustavo Funchal*, Tiago Pedrosa†, Marcos Vallim*, Paulo Leitao†
* Federal University of Technology – Parana (UTFPR)
Email: {gustavofunchall, mvallim}@gmail.com
† Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança,
Campus de Santa Apolónia, 5300-253 Bragança, Portugal
Email: {pedrosa, pleitao}@ipb.pt

2020 IEEE 18th International Conference on Industrial Informatics (INDIN) | 978-1-7281-4964-6/20/$31.00 ©2020 IEEE | DOI: 10.1109/INDIN45582.2020.9478915

***Abstract*—One main foundation of Industry 4.0 is the connectivity of devices and systems using Internet of Things (IoT) technologies, where Cyber-physical systems (CPS) act as the backbone infrastructure based on distributed and decentralized structures. This approach provides significant benefits, namely improved performance, responsiveness and reconfigurability, but also brings some problems in terms of security, as the devices and systems become vulnerable to cyberattacks. This paper describes the implementation of several mechanisms to increase the security in a self-organized cyber-physical conveyor system, based on multi-agent systems (MAS) and build up with different individual modular and intelligent conveyor modules. For this purpose, the JADE-S add-on is used to enforce more security controls, also an Intrusion Detection System (IDS) is created supported by Machine Learning (ML) techniques that analyses the communication between agents, enabling to monitor and analyse the events that occur in the system, extracting signs of intrusions, together they contribute to mitigate cyberattacks.**

***Index Terms*—Multi-agent systems, Cyber-physical systems, Cybersecurity, Machine Learning, Intrusion Detection Systems.**

I. Introduction

Industry 4.0 is driven by emergent technologies that cause profound changes in production systems and business models. This revolution promotes the connectivity of devices and systems through the use of Internet of Things (IoT) technologies, complemented with the use of other disruptive technologies, e.g., Big Data, Cloud and Artificial Intelligence (AI).

Cyber-Physical Systems (CPS) are a backbone approach to develop such smart Industry 4.0 compliant solutions following a decentralized structure [1], [2], based on a network of cyber and physical entities that are integrated to realize a particular objective. CPS can be seen as a way to transform traditional factories into intelligent factories in the Industry 4.0 context [3]. CPS is particularly involved in integrating the paramount of sensors and actuators nodes to achieve certain system goals, and for that, IoT technologies are used to reach collaborative work on distributed systems [4].

Multi-Agent Systems (MAS) is a suitable approach to realize CPS due to its inherent capabilities to distribute intelligence using decentralized structures. In fact, MAS allows the development of large-scale complex engineering problems by decentralizing the control functions over distributed and intelligent software agents, running in cloud and edge computational layers, that cooperate together to achieve the system goals [5]. In such decentralized systems, as MAS are, the data privacy and integrity assumes a crucial issue, since security vulnerabilities appears.

According to the McKinsey report [6], cyberrisk is the biggest threat to business considered by risk managers, and 75% of experts consider cybersecurity a top priority. In addition, only 16% of executives at large companies say their companies are prepared to deal with cyberrisks and this is due to the growth of new technologies, such as IoT and AI, that bring many benefits, but exposes companies and their customers to new cyberrisks that arrive in new ways. Some companies are investing up to USD 500 million on cybersecurity due to the growth of the threat from cyberattacks. A study by Cisco [7] showed that 53% of the attacks resulted in losses of more than USD 500.000, which includes lost revenue, opportunities, customers and out-of-pocket costs.

In this context, the implementation of security mechanisms is crucial to secure these systems, which can range from simple to more complex ones. Simple mechanisms include authentication, signature and encryption, as shown by [8] in the application of these mechanisms in MAS solutions based on JADE (Java Agent Development Framework). The complex mechanisms can be accomplished with the use of ML algorithms that are capable of find patterns in the data and classify them in different categories, e.g., using ML to detect the occurrence of a cyberattack at the physical process layer and also to identify the specific type of the attack [9].

Having this in mind, this paper discusses the benefits of using ML techniques to implement secure industrial CPS based on MAS. Fostering the creation of an Intrusion Detection System (IDS), ML techniques are used both to analyse the messages exchanged between the agents, to identify whether they follow normal patterns or not, as well as to assist in the observation of each message exchanged between the agents, aiming to classify these messages as reliable or not. In addition, other security mechanisms, e.g., user authentication, signature and encryption, have been implemented, covering a significant number of attack vectors to be mitigated.

The remaining paper is organized as follows: Section II presents the security requirements in industrial CPS, particularly those based on MAS. Section III describes the experimental case study related to the cyber-physical conveyor system. Section IV presents the developed cyber-security mechanisms using ML techniques and Section V discusses the achieved results. Finally, Section VI rounds up the paper with the conclusions and points out some future work.

II. Security Issues in CPS

There were no major concerns about the security of control networks in the past, however, with the evolution of internet-based communication, attacks on industrial networks have become increasingly common nowadays. A pertinent issue is related to the security of Networked Control Systems (NCS) concerning Information Technology (IT) security and network security solutions. It is not so simple to deal only with existing solutions, since NCS applications are coupled with physical environments, requiring new security mechanisms [10]. The key parameters that differentiate between IT and NCS security are performance and reliability requirements, operating systems and applications, risk management goals, security architecture, security goals and assumptions. The main goal of IT security is related to data protection while the main goal of NCS security is related to the protection of the process, i.e. maintain the availability and correct execution of the process are the priorities.

In such environments, known threats that can target smart devices include distributed denial of service (DDoS), botnets and malware, Man-in-the-Middle (MitM), data breaches and perimeter attacks that connect IoT devices [11], [12]. These threats can cause serious problems, being able to capture the system information and to make changes in order to interfere with the quality of service, with the command and control of the system devices in an undesired way, which enables the execution of arbitrary codes that can damage system devices or even to shut down the service provided by the device.

Agent-based systems are very similar to smart devices, as MAS distributes intelligence through a society of cooperative and autonomous agents, which allows the design of complex and large-scale systems that are characterized by modularity, robustness, flexibility and responsiveness, being a suitable technology to perform CPS [13]. Since these systems are based on distribution and collaboration, they present security constraints. Some of the attacks on industrial agents are [14]:

- Masquerading: the acquisition of internal information by external entities by faking the identity of trusted entities.
- Eavesdropping: the monitoring of internal and external communication, enabling the direct access to agent operations as well as their data.
- Cloning or Replacement: the agent's behavior is cloned and replaced by a malicious one, allowing to collect system data and execute malicious commands.
- Agent manipulation: the manipulation of the agent states and data to guide their behavior.

Knowing that the systems can be subject to various attacks, the security analysis in the systems' architecture plays an important role in addressing the security threats it contains. The purpose of the threat analysis is to identify, prioritize and mitigate potential security threats, eliminating possible vulnerabilities in the system architecture, which consequently reduces waste in the process and the attack vector.

The JADE-S add-on provides security in MAS solutions developed using the JADE framework, through the introduction of security in the actions performed on the platform and also in the communication between agents through authentication and encryption mechanisms [15]. However, some additional techniques are being used to supervise communication to detect intrusion, such as hybrid deep learning [10], ML [16] or multi agent collaboration models [17]. This reinforces the idea that the use of ML techniques can be important to improve security in agent-based systems, as these techniques are able to introduce more intelligence in agents allowing to identify the malicious behavior of external agents or threats [15].

III. Experimental Case Study

This section aims to describe the experimental case study as well as the security methodology.

A. Description of the Case Study

The case study consists of a conveyor transfer system that is build up of a sequence of Fischertechnik conveyors, where each one transfers parts from an initial to a final position. Each conveyor contains a belt operated by a 24V DC motor and two light sensors responsible for detecting parts in the initial and final positions, being also operated in 24V, as depicted in Fig. 1. As described in [18], this system supports scalability and on-the-fly reconfigurability, i.e. it is not necessary to stop, restart or reprogram the system in cases of adding or removing a conveyor, or even changing the conveyors positions.

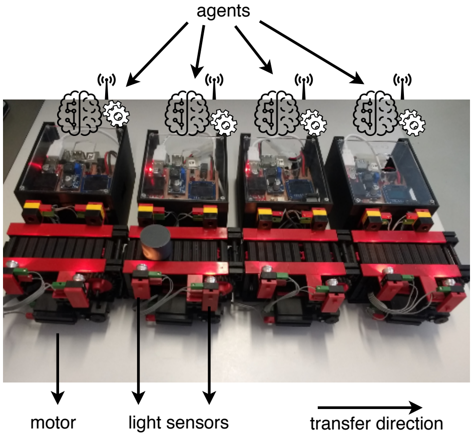


Fig. 1. Modular and self-organized conveyor transfer system.

The on-the-fly reconfiguration approach was performed by using the MAS technology to create a CPS based on several modules of cyber-physical devices, where a software agent, enhanced with self-organization mechanisms, plays the role of the cyber part controlling the physical conveyor device.

Therefore, the cyber-physical conveyor system comprises several modular cyber-physical conveyor devices, each one performing its own operation and interacting with the others to ensure the emergence of the system transfer functionality. The cyber part of each module consists of a Raspberry Pi single board computer, where each agent is running.

The communication between the agents uses the Raspberry Pi's WiFi capability and the messages are formatted according to the FIPA Agent Communication Language (ACL). The number of exchanged messages is strongly dependent of the number of conveyors, since they need to exchange information during the transport of the parts and also to maintain updated their current position in the conveyor system. For this purpose, if there is an addition or removal of a module, or even a change of position between modules, they must recognize the change as soon as possible in order to transfer the part correctly.

In this context, several interaction patterns were designed to exchange information with the adjacent agents regarding the position of the parts. Examples of exchanged messages are: *informIAmAlive*, *tokenTransmissionInput*, *tokenTransmissionOutput*, *thereIsASwap*, *swapTokenFound*, *goodbye*, *pieceAtLastConveyor*, *conveyorCountWarning* and *productIDTransmission*.

B. Security Methodology

The control system depends on the messages exchanged by the agents through the network, presenting a similar architecture to NCS, as depicted in Fig. 2.

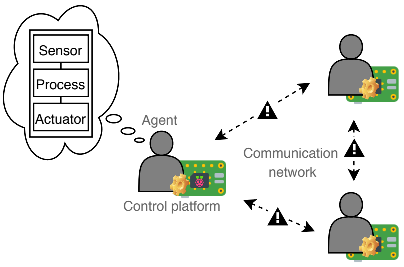


Fig. 2. Control system architecture.

Having this in mind, the vulnerabilities were analyzed in relation to the exchanged messages, by applying some security mechanisms, e.g., encryption, signature, user authentication and ML techniques, to mitigate possible threats. This allows to increase the security in MAS based systems, prioritizing the process protection and system availability. The threat modeling was created, based on the STRIDE methodology [19], to identify the system's vulnerabilities. The mechanisms used in this work have as main focus the prevention of Spoofing, through the user authentication, Tampering, through the verification and analysis of all messages received by the agents using ML techniques, Repudiation, through the message signing, and Information Disclosure, through the message encryption.

IV. Deployed Security Mechanisms in MAS

This section describes the implementation of ML techniques in the described multi-agent cyber-physical conveyor system, and the implementation of the JADE-S add-on features.

A. Methodology

All events that occur in the system, i.e. detection of parts in the input and output light sensors, are perceptions that are captured by the agents, allowing them to make decisions to generate actions, i.e. turning on or off the motor. As it is a distributed and collaborative system, in addition to having perceptions about the environment, the agent must also communicate with other agents to pass its perceptions to others. Therefore, the functioning of the case study system depends mainly on the exchanged messages between the agents, to determine their position in the sequence and to transfer parts.

The messages carry important information that is capable of generating actions and, a non-authorised change of information in its content can cause serious problems in the functioning of the system, e.g., blocking the transport or damaging the system components or the parts being transported.

In this work, the attacks were created in order to explore the attack vectors described in the STRIDE methodology. Therefore, the created hacker agent simulates the attacker's behavior when sending messages exploiting the system's vulnerabilities. These messages are intended to simulate predictable attacks that could directly affect the system. The hacker agent sends malicious messages to the system agents, i.e. messages with non-valid contents that will interfere with the system functioning. For this purpose, the hacker agent filters the gathered messages through message templates and replicates them with content changed.

One of the attacks consists of listening the exchanged messages and whenever a message regarding the *tokenTransmissionOutput* protocol is received, the malicious agent replicates that message with a modified content. This action leads the agents to take actions, e.g turning on the modules' motors, in improper moments, causing several modules to have the motors running unnecessarily as there is no part to be transported at that time. Another attack also implemented in this work is similar to the previous one, but now is the content of the *tokenTransmissionInput* message that is modified, causing the module that is waiting to receive a part to turn off the motor, blocking the transfer of parts throughout the system.

Having created these attacks, all interactions between agents were recorded in a database, enabling the data collection during the normal system operation and also during the attacks. These data were used to create two sets of data that can be used by ML algorithms for two different situations: one for the analysis of the arrival of new messages to the agents and the other for the detection of intrusion to the system.

Having this in mind, the main objective of applying security mechanisms in this system is to protect data and maintain the system functionality and availability. For that, simple mechanisms provided by the JADE-S add-on [20] were implemented on the JADE platform to enable the integrity and confidentiality of messages and ML techniques were deployed in agents to be able to use past experience to predict future decisions, which would make agents made better decisions regarding incoming message patterns.

B. JADE-S Add-on

To protect the exchanged data, the features presented in the JADE-S add-on [20] were used to increase the security of the system. It provides support for security in MAS, e.g., end-to-end message signature and encryption, user authentication and agent action authorization against agent permissions. For the case of user authentication, the special login module was used, which allows a basic authentication using the plain text password file held in a system folder that aims to be used for initial tests. However, standard modules are also available like the Unix, NT and Kerberos, which are easy to apply.

The signature and encryption of messages were applied to the entire payload to protect all information, e.g., content, protocol and ontology. Note that when using these features, users do not have to deal with the signing and encryption mechanisms (e.g., keys and algorithms), and it is only necessary to require that the message be signed and encrypted.

C. Analysis of the Arrival of New Messages

Regarding the implementation of ML techniques, the first approach consists in deploying a ML algorithm in each agent of the system, running in parallel with the defined behavior of the agent, as illustrated in Fig. 3, with the objective to analyze each received message and classify that message as reliable or not. This allows to enhance the agents with a certain intelligence for a better decision-making process.

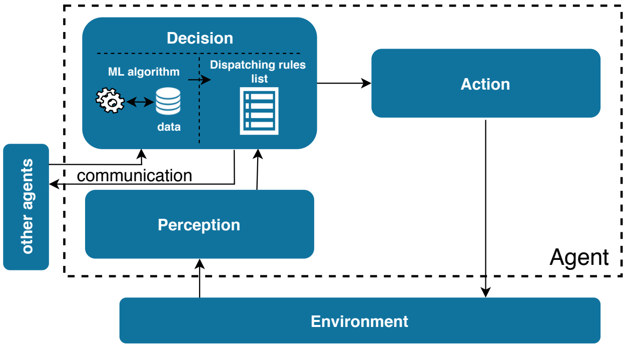


Fig. 3. ML techniques for security incorporated in the agent's behaviour.

The information extracted from the received messages contains the receiver, message content (MC), protocol (P), conveyor identification (CID) in the sequence, message processing time (Time), number of conveyors (NoC), number of messages (NoM) received by the agent in the transport of a part and the classification as reliable or unreliable, as shown in Table I.

TABLE I
SAMPLE OF DATA USED IN THE FIRST APPROACH

Receiver	MC	P	CID	Time	NoC	NoM	Class
Agent4	1	TI	2	5	4	2	Reliable
Agent1	1	TI	3	3	4	2	Reliable
Agent3	1	TI	4	2	4	2	Reliable
Agent1	1	TO	3	1	4	3	Reliable
Agent1	2	TO	3	2	4	4	Unreliable

Regarding the data set, 22365 samples were collected. The training and test data sets were created, with 15655 data in the training set and 6710 data in the test set, representing 70% and 30% of the data respectively. The data is not balanced, with 85% of the data belonging to the reliable class and 15% of data to the unreliable class. This is due to the fact that even attacking the system with some external messages, the system exchanges a lot of information to comply with its operation.

The Random Forest, Decision Trees (DT), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost) and Neural Network classification algorithms were tested. The configuration of the parameters for these models was performed by using the GridSearchCV, which implements an exhaustive search on the parameter values specified for an estimator.

D. Analysis for the Intrusion Detection

The second approach consists of collecting data regarding the passage of each part through all conveyor modules so that the interaction patterns between the agents to transport each part are analyzed and classified as normal or abnormal. Thus, when an external message is received, the algorithm must be able to identify if the transport of the previous part did not have the normal system pattern. Consequently, if anomalies in the functioning of the system are detected, this may indicate that there is a possible intrusion into the system.

The information extracted for the second approach are the number of messages (NoM), average time (AT), processing time (PT), number of conveyors (NoC) and number of protocols used (PU) in the transfer of a part, as shown in Table II.

TABLE II
SAMPLE OF DATA USED IN THE SECOND APPROACH

NoM	AT	PT	NoC	PU	Class
6	35	210	2	3	Normal
9	7	65	2	4	Normal
10	104	1041	2	3	Abnormal
18	1	23	3	4	Normal
25	11	286	3	6	Normal
29	3	99	3	4	Abnormal
42	19	831	4	4	Normal
52	13	717	4	6	Normal
62	10	622	4	4	Abnormal

Regarding the data set in the second approach, 522 samples were collected. The training and test data sets were created with 360 in the training set and 162 in the test set, representing 69% and 31% of the data respectively. Like the previous case, the data is not balanced, presenting 81.6% of the data belonging to the normal class and 18.4% to the abnormal class.

The same algorithms tested in the previous approach were used in this approach, and GridSearchCV was used again to find the best configuration parameters for that data set.

V. Results and System Integration

This section aims to present the achieved results, highlighting the most relevant points and presenting the way in which the algorithms were integrated into the system.

A. *Results and Evaluation Method*

To analyze the success rate of the application of ML algorithms in the case study system, the most well-known metrics presented in [21] were used, namely accuracy, precision, recall and F-measure. Accuracy is the percentage of records correctly classified in the test data set, precision is the percentage of records correctly classified as true on the true set, recall (also known as sensitivity) refers to the percentage of records correctly classified as true within the classified set, and F-measure is the harmonic mean of precision and recall. The achieved results are summarised in Tables III and IV.

TABLE III
RESULTS FOR THE ANALYSIS OF THE ARRIVAL OF NEW MESSAGES

Algorithm	Class	Precision	Recall	F-measure	Accuracy
RF	0	0.98	0.99	0.98	97.25
	1	0.93	0.88	0.90	
DT	0	0.98	0.99	0.98	97.30
	1	0.94	0.87	0.91	
SVM	0	0.95	0.99	0.97	94.02
	1	0.90	0.67	0.77	
MLP	0	0.96	0.97	0.97	94.48
	1	0.84	0.77	0.81	
XGBoost	0	0.98	0.99	0.98	97.37
	1	0.95	0.87	0.91	

TABLE IV
RESULTS FOR THE INTRUSION DETECTION ANALYSIS

Algorithm	Class	Precision	Recall	F-measure	Accuracy
RF	0	0.98	1.00	0.99	98.14
	1	1.00	0.86	0.93	
DT	0	0.99	0.99	0.99	98.76
	1	0.95	0.95	0.95	
SVM	0	0.90	1.00	0.95	90.12
	1	1.00	0.27	0.43	
MLP	0	0.93	0.99	0.96	93.20
	1	0.92	0.55	0.69	
XGBoost	0	0.98	1.00	0.99	98.14
	1	1.00	0.86	0.93	

The results show that the different algorithms for both approaches present excellent results, reflected in the high values for *Precision* and *Recall*. However, the XGBoost algorithm was the one that showed the best result for the first approach, which consists of the analysis of the arrival of new messages. It is possible to observe that for the set of messages considered as reliable, 98% will be classified as reliable and as the sensitivity presents 99%, 1% of these messages may not be really reliable. For unreliable messages, 95% will be correctly classified as unreliable, however 13% of messages in this set can be classified as unreliable and are reliable. For the values of *F-measure*, they present 98% and 90% for classes 0 (reliable) and 1 (unreliable), respectively, indicating that on average, 98% of reliable messages and 91% of messages unreliable will be classified correctly. The accuracy value indicates that the application of this algorithm in the test data set showed 97.37% of correct classifications.

For the second approach, which consists of intrusion detection, the DT algorithm showed the better results. It is possible to observe that for the set of messages considered as normal, 98% will be correctly classified and as the sensitivity presents 100%, these messages will be really normal. Regarding abnormal messages, all messages will be classified correctly, but 14% of messages that are not abnormal will be also considered abnormal. Regarding the *F-measure* values, on average 99% of the normal message sets and 93% of the abnormal message sets will be classified correctly. Finally, this algorithm has 98.76% of correct classifications in the test data set.

In summary, the achieved results can be considered as excellent, since in both approaches high values for precision and recall were observed, which are extremely important to ensure an excellent classification since the data are not balanced. These results showed that the techniques are effective and are able to mitigate problems related mainly to Tampering.

B. *Integration of Algorithms in the System*

The integration of the two algorithms for the different approaches follows a continuous cycle, as depicted in Figure 4. Basically, the second approach (i.e. the intrusion detection), that presented the best results, will be the main technique implemented in the system, being responsible for always analyzing the exchanged messages regarding the transport of a part. If abnormal messages are detected, the first approach (i.e. analysis of the arrival of new messages) is activated so that each message is analyzed separately to mitigate the damage caused to the system. Also, if the second approach detects normal conditions again after a predefined time, the first approach will be disabled.

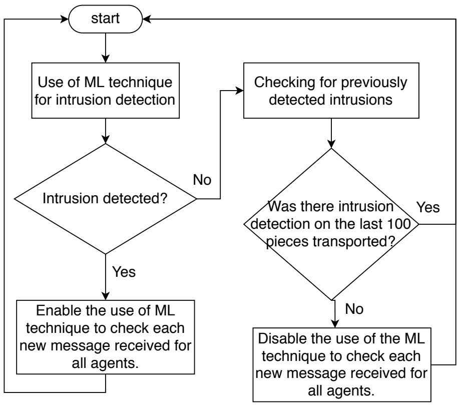
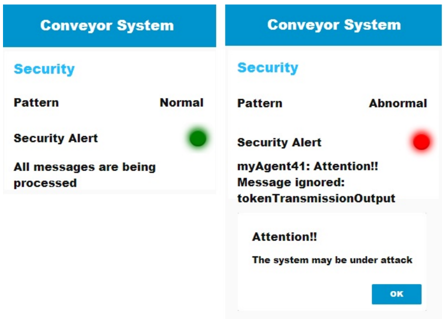

start
Use of ML technique for intrusion detection
Intrusion detected?
No
Yes
Enable the use of ML technique to check each new message received for all agents.
Checking for previously detected intrusions
Was there intrusion detection on the last 100 pieces transported?
Yes
No
Disable the use of the ML technique to check each new message received for all agents.

Fig. 4. Cyclic operation process of the two ML approaches.

The intrusion detection technique is only used by the agent that is the last in the conveyor sequence, but is implemented in all agents to maintain the dynamic functioning and reconfiguration of the system. The technique for classifying the arrival of new messages is implemented and used by all agents.

Besides the deployed ML algorithms, additional rules were implemented in the behavior of each agent, e.g., the verification of the message's sender, accepting only messages sent by known agents belonging to the system. Thus, the approach using ML algorithms is not trained with the sender's data, to be able to recognize abnormal patterns coming from known agents, which favors the identification of messages received by a malicious entity with an ID known by the system. Therefore, this approach is used to protect the exchange data at a low level, when the attackers are able to falsify its identity, e.g., IP address, as well as to capture the data exchanged by agents.

In addition, a graphical interface was created to real-time monitor the security conditions of the system, as depicted in Fig. 5. The user can see the current system condition (Normal/Abnormal) and monitor whether all messages are being processed (if in normal condition) or which messages are being ignored (if in abnormal condition). Whenever the system detects abnormal conditions, a pop-up window is displayed, alerting about a possible attack or abnormal condition.



Conveyor System

Security

Pattern Normal

Security Alert ●

All messages are being processed

Conveyor System

Security

Pattern Abnormal

Security Alert ●

myAgent41: Attention!! Message ignored: tokenTransmissionOutput

Attention!!

The system may be under attack

Fig. 5. Graphical interface to monitor the security conditions.

VI. Conclusions and Future Work

This paper discusses the importance of security issues in industrial CPS based on MAS, primarily aiming to ensure the system availability and the process protection. It was considered a case study related to an automated conveyor transfer system, developed using MAS technology. Vulnerabilities were identified in this system, mainly in terms of eavesdropping and information change, which can interfere with the quality of service.

Having identified the security system weaknesses, some solutions have been implemented to mitigate potential security threats, e.g., the application of the JADE-S add-on for encryption and signature of messages exchanged between agents. More complex and sophisticated mechanisms were deployed using ML techniques to increase the intelligence of the agents in relation to the incoming messages, being able to classify them as unreliable messages. For the classification of messages as reliable or unreliable, the XGBoost algorithm was used obtaining 97.37% of accuracy, while for the intrusion detection, the DT algorithm presented 98.76% of accuracy.

Despite the fact that the attacks implemented in this work are relatively simple, the main objective was to show the application of ML techniques deployed in agents to bring security to the MAS CPS. Future work will be devoted to extend these attacks with more complex scenarios, as well as implement more security ML based mechanisms, mainly aimed at preventing attacks such as Denial of Service.

Acknowledgment

This work has been supported by FCT - Fundação para a Ciência e Tecnologia within the Project Scope: UIDB/05757/2020. This work has been also partially supported by the European Regional Development Fund (ERDF) through the Interreg Spain-Portugal V-A Program (POCTEP) under grant 0677_DISRUPTIVE_2_E (Intensifying the activity of Digital Innovation Hubs within the PocTep region to boost the development of disruptive and last generation ICTs through cross-border cooperation).

References

[1] E. Lee, "Cyber physical systems: design challenges," in *Proceedings of the 11th IEEE International Symposium on Object/component/service-oriented real-time Distributed Computing*, may 2008, pp. 440–451.

[2] P. Leitão, A. Colombo, and S. Karnouskos, "Industrial automation based on cyber-physical systems technologies: Prototype implementations and challenges," *Computers in Industry*, vol. 81, 09 2015.

[3] A. Hellinger, H. S. Translation, J. Macfarlane, B. Services, and H. Galloway, "Cyber-physical Systems Driving Force for Innovation in Mobility, Health, Energy and Production Acatech (ed.)."

[4] A. B. Chebudie, R. Minerva, and D. Rotondi, *Towards a definition of the Internet of Things (IoT)*. IEEE, 2015.

[5] J. Barbosa, P. Leitão, and J. Teixeira, "Empowering a Cyber-Physical System for a Modular Conveyor System with Self-organization," in *Service Orientation in Holonic and Multi-Agent Manufacturing, Studies in Computational Intelligence*. Springer, 2017, vol. 762, pp. 157–170.

[6] T. Poppensieker and R. Riemenschnitter, "A new posture for cybersecurity in a networked world," in *Perspectives on transforming cybersecurity*. McKinsey&Company, 2019, ch. 2.

[7] Cisco, *2018 Annual Cybersecurity Report*, 2018.

[8] X. Vila, A. Schuster, and A. Riera, "Security for a multi-agent system based on JADE," *Computers & Security*, vol. 26, no. 5, pp. 391–400, 2007.

[9] K. Junejo and J. Goh, "Behaviour-based attack detection and classification in cyber physical systems using machine learning," 05 2016.

[10] S. Potluri, N. F. Henry, and C. Diedrich, "Evaluation of hybrid deep learning techniques for ensuring security in networked control systems," in *Proc. IEEE ETFA'17*, 2017, pp. 1–8.

[11] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "DDoS in the IoT: Mirai and Other Botnets," *Computer*, vol. 50, no. 7, pp. 80–84, 2017.

[12] P. Varga, S. Plosz, G. Soos, and C. Hegedus, "Security Threats and Issues in Automation IoT," in *2017 IEEE 13th International Workshop on Factory Communication Systems (WFCS)*, May 2017, pp. 1–6.

[13] P. Leitão, S. Karnouskos, L. Ribeiro, J. Lee, T. Strasser, and A. W. Colombo, "Smart agents in industrial cyber–physical systems," *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1086–1101, May 2016.

[14] S. Karnouskos, "Industrial agents cybersecurity," in *Industrial Agents*. Elsevier, 2015, pp. 109–120.

[15] J. D. Morenas, C. Miller, G. Funchal, V. Melo, M. Vallim, and P. Leitão, "Security Experiences in IoT based applications for Building and Factory Automation," in *Proc. IEEE ICIT'20*, 2020, pp. 322–327.

[16] M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan, and R. Jain, "Machine Learning-Based Network Vulnerability Analysis of Industrial Internet of Things," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6822–6834, Aug 2019.

[17] L. Rafferty, F. Iqbal, S. Aleem, Z. Lu, S. Huang, and P. C. K. Hung, "Intelligent Multi-Agent Collaboration Model for Smart Home IoT Security," in *IEEE Int'l Congress on Internet of Things*, 2018, pp. 65–71.

[18] P. Leitão, J. Barbosa, G. Funchal, and V. Melo, "Self-organized Cyber-Physical Conveyor System using Multi-agent Systems," *Accepted to be published in the International Journal of Artificial Intelligence*, 2020.

[19] R. Khan, K. McLaughlin, D. Laverty, and S. Sezer, "STRIDE-based threat modeling for cyber-physical systems," in *IEEE PES Innovative Smart Grid Technologies Conf. Europe (ISGT-Europe)*, 2017.

[20] JADE Board, *JADE Security Guide*. TILAB S.p.A., 2005.

[21] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation," in *AI 2006: Advances in Artificial Intelligence*. Springer, 2006, pp. 1015–1021.