An Artificial Intelligence Cyberattack Detection System to Improve Threat Reaction in e-Health

Carmelo Ardito[2], Tommaso Di Noia[2], Eugenio Di Sciascio[2], Domenico Lofù[1,2], Andrea Pazienza[1] and Felice Vitulano[1]

[1]*Innovation Lab, Exprivia S.p.A. – Via A. Olivetti 11, Molfetta (I-70056), Italy*
[2]*Politecnico di Bari – Via E. Orabona 4, Bari (I-70125), Italy*

Abstract

In the e-Health domain, new and continuously evolving threats emerge every day. The security of e-Health telemonitoring systems is no longer negligible. In this paper, we propose a Cyberattack Detection System (CADS) model that exploits artificial intelligence techniques to detect anomalies without requiring a security analyst, explain the malicious activity, and display suspected attack data to healthcare personnel for feedback. The system description is contextualized to the case of the hacked remote patient health telemonitoring.

1. Introduction

Security plays a major role in the healthcare domain. Preventing cyberattacks on healthcare infrastructures is no longer negligible. Compromising security in any e-Health system can lead to serious damage to patients' health. In particular, in a remote care context, the protection of telemonitoring systems of patients are essential to ensure that they follow their Clinical Pathway without any kind of external intrusion.

Artificial Intelligence (AI) plays an important role in combating cyberattacks on the security of patient telemonitoring systems [1, 2, 3]. A system that monitors and prevents cyberattacks in healthcare not only must detect the attack, but should also be able to properly understand and report it to the user. In particular, Anomaly Detection systems are renowned approaches that are based on Machine Learning (ML) or Deep Learning (DL) methods to model normal activity in a such way as to easily detect abnormal deviations from the standards in a data-driven fashion. Therefore, in such a sensitive domain, where several healthcare professionals are involved, in addition to detecting threats, it is of paramount importance to represent and explain them through appropriate Explainability algorithms [4]. Moreover, current detection models and rules are not mature enough to recognise early breaches that have not yet caused any damage.

Intrusion analysts infer the context of the incident using prior knowledge to discover events

ITASEC21: Italian Conference on Cybersecurity, April 07–09, 2021
carmelo.ardito@poliba.it (C. Ardito); tommaso.dinoia@poliba.it (T. Di Noia); eugenio.disciascio@poliba.it (E. Di Sciascio); domenico.lofu@exprivia.com (D. Lofù); andrea.pazienza@exprivia.com (A. Pazienza); felice.vitulano@exprivia.com (F. Vitulano)
0000-0001-8993-9855 (C. Ardito); 0000-0002-0939-5462 (T. Di Noia); 0000-0002-5484-9945 (E. Di Sciascio); 0000-0001-6413-9886 (D. Lofù); 000-0002-6827-2585 (A. Pazienza); 0000-0002-6059-8177 (F. Vitulano)

relevant to the incident and understand why it happened [5]. Although security tools that provide visualization techniques and minimize human interaction have been developed to make the analysis process easier, too little attention has been given to making human-friendly the interpretation of security incidents. Simply reporting a cyberattack in written format is not enough to enable the healthcare professional to correct the patient's Clinical Pathway. This data must be represented in a graphical way, which can be grasped by the healthcare provider. The detection of the cyberattack must therefore be supported by systems that provide different forms of explanation, depending on the different end users, and that allow these users to have the possibility to interactively manipulate graphical representations based on Visual Data Mining techniques.

The Internet of Things (IoT) has transformed hospital settings and created a new moniker for the healthcare world, The Internet of Medical Things (IoMT). Ensuring a security mechanism for IoMT, which uses appropriate analytical tools in a distributed working architecture, also capable of analyzing huge data (i.e., big data) generated by IoMT devices in a distributed manner, is a challenging issue. In this paper, we present a cyberattack detection model that implements eXplainable AI (XAI) functionality to support caregivers in grasping that an attack has occurred to the telemonitoring system and its effect on the patient's Clinical Pathway. The main contribution is a *Cyberattack Detection System* that is able to identify compromised data during a cyberattack and support the caregiver to fix the problem. In this paper, we focus on the case where an IoMT system has been hacked, with the malicious intent of manipulating health data, in order to trigger bogus care countermeasure.

The remainder of the present work is as follows. Section 2 provides an overview of related work and technologies which were investigated as background knowledge, namely, cyber-attack detection systems, anomaly-based intrusion detection systems, anomaly detection in e-Health and IoMT with AI techniques. Section 3 describes the architecture of our proposal, i.e. a Cyberattack Detection System with an anomaly detector, an explanation module and a user interface engine, showing a possible scenario of application specifically designed for our approach. Finally, Section 4 concludes the paper, outlining future works.

2. Background and Related Work

Cyberattack detection can be defined as the problem of identifying individuals who are using a computer system without authorization, those who have legitimate access to the system but are abusing their privileges, and, in general, the identification of attempts to use a computer system without authorization or to abuse existing privileges.

In this landscape, modern cyberattack detection systems monitor either host computers or network links to capture cyberattack data. Host intrusion detection refers to the class of intrusion detection systems (IDS) that reside on and monitor an individual host machine [6]. There are a number of system characteristics that a host intrusion detection system (HIDS) can make use of in collecting data [7]. A network intrusion detection system (NIDS), instead, monitors the packets that traverse a given network link. Such a system operates by placing the network interface into promiscuous mode, affording it the advantage of being able to monitor an entire network while not divulging its existence to potential attackers [8].

Cyberattack Detection System (CADS) is software that automates the cyberattack detection process and detects possible cyberattacks. Cyberattack Detection Systems serve three essential security functions: they monitor, detect, and react to unauthorized activity by company insiders and outsider cyberattack. One of the major approaches to cyberattack detection is Anomaly Detection. It assumes that a cyberattack will always reflect some deviations from normal patterns. In this sense, Anomaly-based IDS compares a model of normal behavior against the incoming traffic in order to find anomalies [9, 10].

Once an intrusive incident has been reported by means of a correct detection, the reaction phase has to be fired, evaluating the impact of this event on the security level of the system [11]. It must provide the set of countermeasures to quickly eradicate the cyberattack and, at the same time, indicate the set of actions to heal the system and bring it back to its normal state. A possible way to react is via the use of Intrusion Response Systems (IRSs), as they are IDSs capable of counteracting suspicious activities [12]. Although intrusion response components are often integrated with the detection ones, they have received considerably less attention than IDS research.

Anomaly Detection typically operates on monitored networked traffic data. Actually, continuous monitoring is the main activity of modern e-Health technologies, ranging from devices that monitor health and deliver medication, to telemedicine delivering care remotely. Indeed, the integration of healthcare-based devices and sensors within IoT, led to the evolution of IoMT [13]. Therefore, IoMT-enabled devices have made remote monitoring possible in the healthcare sector, enabling the ability to keep patients safe, and inspiring doctor to provide superlative treatment [14]. As a result, the increasing demands and expansion of IoMT systems require advancements in data storage methods, data processing and cybersecurity related issues.

Healthcare providers can then provide efficient remote healthcare communication for monitoring and diagnosis services to the residents of these smart communities. Any security threat to these systems may cause a serious problem, such as imposing a false diagnosis or delaying the interaction. This leads to a violation of patients' privacy, health issues, and even death in extreme cases [15]. AI and Machine Learning (ML) have been largely employed for managing issues in healthcare systems as they are the most promising techniques to be used for previously unseen attacks [16]. It can identify attacks simply by monitoring data alteration or by detecting changes in the network's traffic characteristics. In particular, ML-based anomaly detection systems are crucial to ensure security and mitigate threats such as false data injection attacks [17]. IoMT systems are widely distributed and are collections of heterogeneous sensors. Attack detection in IoMT is entirely different from the present security mechanism, due to the special services required by IoMT such as: computing power, memory space, battery life, low latency, and network bandwidth, which cannot get fulfilled by the centralized conventional approach of standalone cloud computing [18]. In cloud computing architecture, data generated by IoMT devices is being transmitted to and from the cloud in order to provide services to the healthcare users. The limitations of traditional standalone cloud solutions is that the data recovery time is too high for a real-time emergency situation, such as fall detection or stroke prevention, which mostly needs rapid response time from medical professionals [19].

Therefore, designing a distributed security framework, for distributed IoMT applications is a challenging task due to the dynamic nature of IoMT network such as IoT devices, edge devices, and cloud. Moreover, the evolution of attacker behavior can intercept the transmission network

in IoMT [20]. The line of research in this way is going towards constructing robust anomaly-based IDSs that efficiently distinguish attack and normal observations in IoMT environment, consisting of interconnected devices and sensors, with poor design and weak authentication measures. As stated in [21, 22], the collection of remote data from these sensors is a complex process due to the different types of devices that are involved to measure the parameters. For this reason, works in [23, 24] dealt with a clinical and operational context to develop integrated solutions for seamless care in which AI and IoMT are used at the Edge, with a people-centered approach that adapt to the needs of healthcare providers and that are embedded into their workflows. Recently, in [25] proposed an ensemble learning model that combines Decision Trees, Naive Bayes, and Random Forest to feed a final XGBoost classifier in order to identify normal and attack instances in an IoMT network. Also, authors in [26] have designed a a real-time Enhanced Healthcare Monitoring System (EHMS) test-bed that monitors the patients' biometrics and collects network flow metrics. Some recent works [27, 28, 29] have proposed to improve the performance of anomaly detection by incorporating a type of feedback from the user, called User Feedback. Nevertheless, all the studies lack of effective threat reaction phases that can be managed with appropriate explainable modules of ML-based models and user feedback modules to ascertain that a detected anomaly is assumed to be malicious.

3. The Cyberattack Detection System

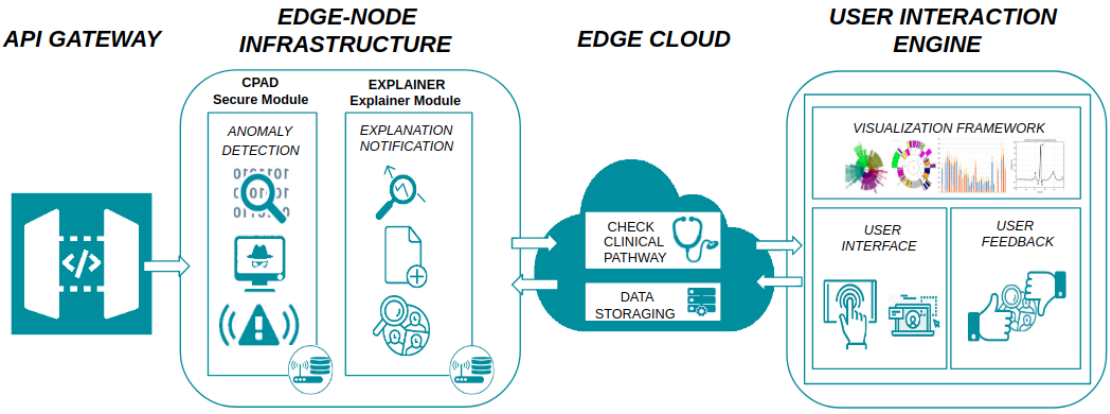

Figure 1: Architecture for the Cyberattack Detection System.

Figure 1 depicts the architecture of the proposed Cyberattack Detection System. An early version was already presented in [30]. This revised and refined architecture focuses on the security of data transmitted from IoMT sensors to three different interconnected processing modules, namely the Clinical Pathway Anomaly Detection (CPAD), the Explainer module, and the User Interaction Engine. The latter is made up of three sub-modules, i.e.: Visualization Framework, User Interface, and User Feedback.

The system implements a methodological approach to the problem of anomaly detection by

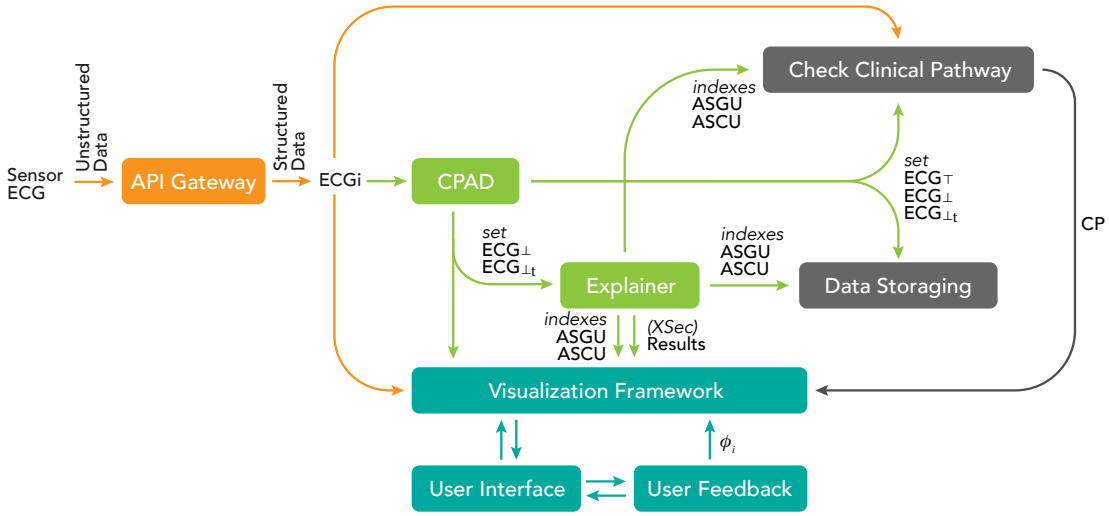

Figure 2: Data flow among system modules.

including, in addition to the identification of anomalous data, an explanation of the possible motivations for classifying such data as anomalous, and the possibility for a domain expert to validate data through their visual interactive representation. Anomalous data, which represent an intrusion in a hacked system, are explained according to the *Explainable Security (XSec)* paradigm [31]. As a result of a hacked home-care telemonitoring system, the Cyberattack Detection System classifies some ECG instances as False Positives (FP) or False Negatives (FN). After the detection, the user will be able to analyze the characteristics of what caused the classification of some data as FP or FN and interact with it. The interaction activity with the data is performed by means of the *User Interaction Engine*, which provides a dashboard through which to visually explore the data to get a clearer view of what happened in a time interval. Furthermore, thanks to the *User Feedback* sub-module, it is possible to implement a continuous improvement of the classification performances, and consequently of the anomaly detection, thus achieving a more robust identification of threats. Figure 2 shows the relationships between the various modules in terms of data flow.

Our system is beneficial in e-Health scenarios, supporting the patient who is in home-based healthcare. The technological infrastructure, based on the architecture presented in [30], promotes the care of a patient according to his or her Clinical Pathway (CP), i.e. a set of diagnostic and therapeutic procedures related to the treatment of that specific patient. The CP represents a tool used to manage the quality in healthcare concerning the standardization of care processes. Its implementation reduces the variability in clinical practice and improves outcomes, aiming at promoting organized and efficient patient care based on evidence-based medicine, and to optimize outcomes in settings such as acute care and home care. A single CP may refer to multiple clinical guidelines on several topics in a well specified context. In this way, some activities can be managed by the health personnel of health structures; some others can be

managed autonomously by the patient, in a sort of medical-unsupervised manner. The home care infrastructure, i.e. the *IoMT-Edge-Computing*, promotes a kind of distributed edge computing of the IoMT network, thus reducing latency and improving reliability. Patient monitoring devices are then connected in the IoMT network. In turn, edge devices communicate with a cloud infrastructure to store gathered clinical data and keep in touch with the corresponding medical staff. Therefore, some vulnerabilities may arise regarding the security of patient and clinical data.

The proposed Cyberattack Detection System implements cybersecurity methods to identify which data were compromised after the system is hacked. Moreover, it provides an explanation of the cyberattack and enables interaction with the detected anomalies which may occur in the remote and continuous patient monitoring and care phases. In particular, the anomaly detection phase is carried out by the CPAD module, whose formalization, already tailored in the healthcare domain, was introduced in [32]. The CPAD detects deviations from the patient's CP and avoids the processing of inconsistent or false data, which could be life-threatening for a patient. After the detection phase, the Explainer module analyzes clinical data classified as anomalous and, through the User Interaction Engine, a validation request is sent remotely to the medical staff. A continuous telemonitoring of patient's clinical parameters is performed, without the need for a physical presence of the health operator. The system, by means of IoMT sensors connected to the Edge network, collects different clinical parameters, such as Electroencephalogram (EEG), Blood Oxygen Level (OXI), Electrocardiogram (ECG), Electromyography (EMG), ALT Blood Test (ALT), and body temperature. Afterwards, such clinical data are processed at the Edge and the patient Clinical Pathway is generated.

In the following, a running example in which an e-Health telemonitoring system has been hacked is reported. It describes how the proposed Cyberattack Detection System is able to highlight the cybersecurity threats and the related countermeasures to solve the hack and restore the system. A male patient who is following a certain CP is monitored. He suffers from Congestive Heart Failure and his CP requires that his heartbeats are monitored every 15 minutes. A smart end-device that measures the ECG is used. Hence, the proposed system, suitably connected to the end-device, receives data about the patient pulse. The gathered heartbeats are fed to the CPAD module that determines whether anomalous measurements occur, suggesting that the telemonitoring system has been hacked. Since heartbeats measures are a key factor in determining the clinical picture of a patient, a compromised flow of measurements would endanger the patient's CP and induce a wrong handling of the patient's health.

The Cyberattack Detection System is designed to be modular and can be integrated with various IoMT devices that use Bluetooth technology. An API gateway is responsible of the correct integration of the end-device with the Cyberattack Detection System. In particular, the API gateway checks the compliance of the gathered data to be fed into the system. By means of specifically designed APIs, input data are normalized according to the system standards, allowing the correct exchange of information between the devices and the software modules, also converting unstructured data into structured ones.

Referring to the running example, in which a smart ECG monitoring end-device is connected via IoMT with the Cyberattack Detection System, the following definition to formally handle data inflow is proposed.

Table 1
ECG device data table format.

ECG	HB_{ts_1}	HB_{ts_2}	$\cdots$	HB_{ts_u}
ECG_0	$x_{(0,1)}$	$x_{(0,2)}$	$\cdots$	$x_{(0,u)}$
ECG_1	$x_{(1,1)}$	$x_{(1,2)}$	$\cdots$	$x_{(1,u)}$
ECG_2	$x_{(2,1)}$	$x_{(2,2)}$	$\cdots$	$x_{(2,u)}$
$\vdots$	$\cdots$	$\cdots$	$\ddots$	$\vdots$
ECG_i	$x_{(i,1)}$	$x_{(i,2)}$	$\cdots$	$x_{(i,u)}$

Definition 1. *Let ECG be the data-flow of a smart ECG monitoring end-device, HB be the heartbeat information coming from ECG at a certain timestamp ts. The i-th heartbeat detection is defined as follows:*

$$ECG_i = f(HB_{ts_u}), \text{with } HB_{ts_u} \in \mathbb{R} \text{ and } u \in [1, l], l \in \mathbb{R} \tag{1}$$

The variable HB_{ts_u} indicates the count-based feature representing the value of the u-th sampling step in a given timestamp, which can be assumed as a real value $x_{(i,u)} \in \mathbb{R}$, representing the amount in milliVolt (mV) of the count-based feature. Therefore, the representation of the ECG data can be formalized as in Table 1.

3.1. ECG Anomaly Detection

The proposed system, through the use of AI techniques, contributes to improve the security of the telemonitoring infrastructure. Once the data have been transformed into a structured form, they are given as input to the CPAD which analyzes the structured data according to Table 1 format, and checks in which point the cyberattack has been launched by retrieving the anomaly. The CPAD module is based on Robust Deep Autoencoders (RDA) [33]. In fact, some recent works demonstrated how these deep approaches perform quite well in detecting cyberattacks by carrying out anomaly detection by means of neural structures [34, 35, 36]. The main advantage of applying RDA for anomaly detection is the capability of discovering high-quality nonlinear features, while at the same time identifying and eliminating outliers and noise.

In the running example, in which it is assumed that a cyberattack on ECG measurements is going on, a delayed threat reaction would compromise the patient's care pathway. By exploiting RDA for anomaly detection, the CPAD modules ensure a rapid response in terms of inference time, quickly detecting anomalous ECG data. In particular, the CPAD module categorizes ECG data with a simple binary classification according to whether they have been identified as anomalous or not. Actually, in the medical literature [37], heartbeats can be classified into five different types:

1. (N) - Normal;
2. (RonT) - Premature Ventricular Contraction;
3. (PVC) - Premature Ventricular Contraction;
4. (SP) - Supra Ventricular;
5. (UB) - Unclassified Beat.

Therefore, the classification process can be further specified by dividing the normal heartbeats from the others, and then dedicate another RDA classifier to detect the remaining four different types of anomalies. In particular, the CPAD module defines three sets of classified ECG data:

- $ECG_{\top}$: instances classified by CPAD as normal, i.e. not anomalous;
- $ECG_{\bot}$: instances classified by CPAD as anomalous;
- $ECG_{\bot_t}$: instances of anomalous ECG, with $t \in \{$(RonT), (PVC), (SP), (UB)$\}$.

Once the structured data have been processed, the CPAD produces a CSV file containing all the instances classified by CPAD as anomalous $ECG_{\bot}$, i.e. all the heartbeats that are anomalous. A confusion matrix is also showed, in order to have a clearer view of performance classification, both in terms of predictability power and effectiveness of the learned model. To address the anomaly detection problem, the RDA method based on the autoencoder approach is exploited. It is based on a training pipeline, where examples of anomalies are provided during model training. The result is evaluated on sets consisting of anomalous and normal (i.e., non-anomalous) data. In the end, the CPAD module separates the predicted instances into the three different sets of heartbeat anomalies, namely $ECG_{\top}$, $ECG_{\bot}$, and $ECG_{\bot_t}$.

3.2. ECG Explanation

After the anomaly detection, the next step is to interpret and explain the sets that CPAD has created, to understand why certain instances have been classified as anomalous. In systems using Explainable AI (XAI) algorithms, additional data coming from the Machine Learning process may be useful explanations, produced by the system itself to enrich the predicted instances with a plausible rationale. In Information Security, instead, explanations are provided by the designers. However, the role of Explanations is crucial in AI field. Consequently, Information Security kept all advantages from that, for example, ensure the user trust concerning the system. Explanations are therefore designed to bridge the gap between "actual safety" and "perceived safety".

To this end, the Explainer module of our Cyberattack Detection System receives as input the i-th ECG instance (ECG_i) from the API Gateway, and also the output of the CPAD, appropriately separated in the three sets $ECG_{\top}$, $ECG_{\bot}$, and $ECG_{\bot_t}$. Exploiting the XSec paradigm, the aim is to obtain its "*six W*" (Who? What? Where? When? Why? and How?) that give a complete view of the identified and perceived anomaly. XSec involves several actors (e.g., in this case we have security analysts, doctors, nurses, and the patient). It requires a dedicated reasoning tool to infer about the system model, the threat model, and properties of security, privacy, and trust, as well as concrete cyberattacks, vulnerabilities, and countermeasures.

In addition to the XSec paradigm, two specifically designed indexes, called respectively *Anomaly Score General Unsafe* (ASGU) and *Anomaly Score Class Unsafe* (ASCU), are introduced. First, we formally define the ASGU index as follows.

Definition 2. *Let ECG be the set of heartbeat detections received from a smart ECG monitoring end-device, ECG_i be the i-th heartbeat detection, and $ECG_{\bot}$ the set of all anomalous instances. Then, the Anomaly Score General Unsafe index is a function* $\mathrm{ASGU}\colon ECG \mapsto [0, 2]$ *which gives a general ranking of possibility of considering a heartbeat as anomalous, according to instances*

in $ECG_{\perp}$. Therefore, $\mathrm{ASGU}(ECG_i) \in \mathbb{R}^{[0,2]}$ indicates how the instance ECG_i is currently considered an anomaly over the whole set $ECG_{\perp}$.

Hence, the higher the ASGU score value, and thus closer to 2, the more anomalous the heartbeats will be considered. The ASCU index, instead, is based on the comparison with one of the four classes of anomalies.

Definition 3. *Let ECG be the set of heartbeat detections received from a smart ECG monitoring end-device, ECG_i be the i-th heartbeat detection, and $ECG_{\perp_t}$ a set of a class of anomalous instances, with $t \in \{$(RonT), (PVC), (SP), (UB)$\}$. Then, the Anomaly Score Class Unsafe index is a function $\mathrm{ASCU}\colon ECG \mapsto [0,1]$ which gives a general ranking of probability of considering a heartbeat as anomalous, according to a class $ECG_{\perp_t}$. Therefore, $\mathrm{ASCU}(ECG_i) \in \mathbb{R}^{[0,1]}$ indicates how the instance ECG_i is currently considered an anomaly over the whole set $ECG_{\perp_t}$.*

In this case, the higher the value of the ASCU score, and therefore the closer to 1, the more abnormal the heartbeats will be in the set of heartbeats of the same class $ECG_{\perp_t}$. The goal of these two indexes is to quantify the strength of a feature in contributing to determine an anomaly in the ECG data. This would give a further useful information to characterizing the detected anomaly, and such information can be assumed as part of the explanation to be addressed together with the XSec approach.

In the running example, the Explaination module receives as input the CSV file containing the ECG_i instances and provides in output another CSV file composed of four columns:

- ECG_i: i-th instance classified by the CPAD as anomalous;
- CLASS: the corresponding class $t \in \{ECG_{\perp_t}\}$ of the i-th instance;
- ASGU: the Anomaly Score General Unsafe score;
- ASCU: Anomaly Score Class Unsafe score.

The Explanation module, therefore, acts in two conjunct phases: the first one takes place following the generation of the CSV file containing the information and indexes mentioned above (ECG_i, CLASS, ASGU and ASCU); the second one results from the information generated by the XSec approach, which will be displayed in the Visualization Framework. Thanks to the combination of the XSec paradigm and the two indexes ASGU and ASCU, the Cyberattack Detection System provides the doctor with a concise explanation of why the i-th detection has been classified as anomalous.

Following the Explanation, it is possible to check whether the Clinical Pathway that has been generated is correct or not. An incorrect CP has a double meaning: from a clinical point of view, it is a serious problem for the patient's health while, from the cybersecurity point of view, it means that the system has been hacked.

3.3. ECG User Interaction

In our case study, the contribution of the User Feedback to the system is the evaluation of a the detected anomaly and the embedding of a doctor's feedback. The User Feedback module will generate for each detection ECG_i a feedback coefficient ϕ_i that represents the doctor's feedback on a given instance.

Definition 4. *Let ECG be the set of the heartbeat detections received from a smart ECG monitoring end-device, ECG_i be the i-th heartbeat detection. Then, the feedback coefficient is a function $\phi\colon ECG \mapsto \{-1, 1\}$ such that any i-th user feedback related to the heartbeat detection ECG_i, is defined as follows:*

$$\phi_i = \begin{cases} +1 & \text{if } ECG_i \text{ is false positive or false negative} \\ -1 & \text{if } ECG_i \text{ is true positive or true negative} \end{cases} \tag{2}$$

Therefore, the User Feedback UF is a set of tuples such that, for any i-th pair of arguments (ECG_i, ϕ_i), a single element UF_i is defined as:

$$UF_i = (ECG_i, \phi_i) \tag{3}$$

In this way, the Cyberattack Detection System will become more robust to external cyberattacks, since the User Feedback would report the opinion of the caregiver which will confirm or not whether the i-the ECG detection is abnormal or not. In the User Interaction Engine, the *Visualisation Framework* represents the data orchestrator, handling and visualizing processed data coming from the various modules. It uses algorithms of Visual Data Mining (VDM) [38, 39] that allow, through different visualisation techniques, to interactively group data in a more efficient way, improving the data insight process.

Afterwards, the *User Interface (UI)* included in the User Interaction Engine allows the user to interact with the data. In the running example, the UI allows the caregiver to interact with the ECG instances. After the Explanation module has displayed the result of the *XSec* paradigm, it is possible to visually manage each ECG detection. For instance, one would be able to no longer consider an ECG instance as an anomaly, or, more specifically, to improve the classifier performances by indicating the correct class of anomaly among the four types $ECG_{\perp_t}$ when a wrong one has been predicted. The interaction with the user, in this case a doctor, helps the system to be more and more reliable, as well as secure from cyberattacks.

Through the integration of CPAD, Explainer, and User Engine Interface modules, the Visualization Framework will be able to manage anomalies detected as *Threat Insight*. These will be appropriately displayed on the UI which, in addition to allowing interaction with the anomalous data (in this case the ECG detection), will be able to display the threat representation through a dashboard. Thanks to the threats graphical representation in the dashboard, the user's reaction to the threat is improved.

4. Discussion and Conclusion

In the e-Health domain, new and continuous evolving threats emerge every day. The security of e-Health telemonitoring systems is no longer a negligible task. The proposed Cyberattack Detection System, thanks to the use of AI techniques, is able to protect the e-Health system from cyberattacks by automatically identifying the anomalies in the e-Health system, without the need of a dedicated security analyst.

The solution is focused on the task of cyberattack detection, in the particular case of exploiting a remote patient telemonitoring system that has been hacked. A specific running example, i.e. the heartbeat telemonitoring, has been considered.

The presented system is designed to automatically detect the anomaly by means of Deep Learning techniques. In particular, a Robust Deep Autoencoder detects anomalous heartbeats instances. The detected anomalous heartbeats are subsequently interpreted with a combination of state-of-the-art explainable security paradigms (XSec) and with two new explainable scores which have been introduced, showing to the user the reasons of a malicious activity interfering with the heartbeats telemonitoring. Further, the systems visually represent the results to the user who is engaged in a feedback response: if the detected anomalous data are truly atypical, the user assess the detection with a positive feedback coefficient, i.e. a new metric involved in the process of user interaction with the system. Otherwise, the user provides a negative feedback coefficient. Such an interaction with the Cyberattack Detection System proposed has an impact on the processed health data, adjusting the visualization reports with the corrected measurements, shown in a useful dashboard.

The proposed work witnesses how the wide range of AI techniques are already dramatically important in the hot topic of Cybersecurity, in particular when applied to eHealth domain. The effectiveness of AI algorithms in identifying threats, and the consequent efficiency in significantly reducing the user's reaction time, are becoming fundamental properties to be taken into account when designing a Cyberattack Detection System. For these reasons, Threat Intelligence is increasingly becoming a tool for enhancing system security and performing automatic threat analysis. Indeed, Threat Intelligence uses different data sources (in this case data from different IoMT sensors) to proactively identify and mitigate threats to improve the security of the patient telemonitoring system. Overall, the presented system suggests how it is possible to improve user reaction to threats in healthcare telemonitoring systems using AI techniques.

Future work will be devoted to the integration of other IoMT sensors and, hence, to the anomaly detection of other vital parameters.

Acknowledgments

This work was partially funded by the European Union, Horizon 2020 research and innovation programme, through the ECHO project (grant agreement no 830943) and by the Italian P.O. Puglia FESR 2014 – 2020 (project code 6ESURE5) SECURE SAFE APULIA.

References

[1] S. Bhaskar, S. Bradley, S. Sakhamuri, S. Moguilner, V. K. Chattu, S. Pandya, S. Schroeder, D. Ray, M. Banach, Designing futuristic telemedicine using artificial intelligence and robotics in the covid-19 era, Frontiers in public health 8 (2020) 708.

[2] R. D. Kindle, O. Badawi, L. A. Celi, S. Sturland, Intensive care unit telemedicine in the era of big data, artificial intelligence, and computer clinical decision support systems, Critical care clinics 35 (2019) 483–495.

[3] D. M. M. Pacis, E. D. Subido Jr, N. T. Bugtai, Trends in telemedicine utilizing artificial intelligence, in: AIP conference proceedings, volume 1933, AIP Publishing LLC, 2018, p. 040009.

[4] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, H. Müller, Causability and explainability of artificial intelligence in medicine, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 9 (2019) e1312.
[5] N. AfzaliSeresht, Q. Liu, Y. Miao, An explainable intelligence model for security event analysis, in: Australasian Joint Conference on Artificial Intelligence, Springer, 2019, pp. 315–327.
[6] S. Singh, S. Silakari, A survey of cyber attack detection systems, International Journal of Computer Science and Network Security 9 (2009) 1–10.
[7] J. Raiyn, et al., A survey of cyber attack detection strategies, International Journal of Security and Its Applications 8 (2014) 247–256.
[8] B. Mukherjee, L. T. Heberlein, K. N. Levitt, Network intrusion detection, IEEE network 8 (1994) 26–41.
[9] A. C. Kim, W. H. Park, D. H. Lee, A study on the live forensic techniques for anomaly detection in user terminals, International Journal of Security and Its Applications 7 (2013) 181–188.
[10] D. J. Weller-Fahy, B. J. Borghetti, A. A. Sodemann, A survey of distance and similarity measures used within network intrusion anomaly detection, IEEE Communications Surveys & Tutorials 17 (2014) 70–91.
[11] M. Gunasekharan, S. Basu, G. R. Santhanam, Selecting the minimal set of preferred responses to counter detected intrusions, in: Proceedings of the 12th Annual Conference on Cyber and Information Security Research, 2017, pp. 1–8.
[12] S. A. Zonouz, H. Khurana, W. H. Sanders, T. M. Yardley, Rre: A game-theoretic intrusion response and recovery engine, IEEE Transactions on Parallel and Distributed Systems 25 (2013) 395–406.
[13] G. Yang, M. A. Jan, V. G. Menon, P. Shynu, M. M. Aimal, M. D. Alshehri, A centralized cluster-based hierarchical approach for green communication in a smart healthcare system, IEEE Access 8 (2020) 101464–101475.
[14] H. Zhu, C. K. Wu, C. H. Koo, Y. T. Tsang, Y. Liu, H. R. Chi, K.-F. Tsang, Smart healthcare in the era of internet-of-things, IEEE Consumer Electronics Magazine 8 (2019) 26–30.
[15] H. Fotouhi, A. Causevic, K. Lundqvist, M. Björkman, Communication and security in health monitoring systems–a review, in: 2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC), volume 1, IEEE, 2016, pp. 545–554.
[16] A. L. Buczak, E. Guven, A survey of data mining and machine learning methods for cyber security intrusion detection, IEEE Communications surveys & tutorials 18 (2015) 1153–1176.
[17] Y. Luo, Y. Xiao, L. Cheng, G. Peng, D. D. Yao, Deep learning-based anomaly detection in cyber-physical systems: Progress and opportunities, arXiv preprint arXiv:2003.13213 (2020).
[18] A. Alrawais, A. Alhothaily, C. Hu, X. Cheng, Fog computing for the internet of things: Security and privacy issues, IEEE Internet Computing 21 (2017) 34–42.
[19] S. Singh, Y.-S. Jeong, J. H. Park, A survey on cloud computing security: Issues, threats, and solutions, Journal of Network and Computer Applications 75 (2016) 200–222.
[20] R. Priyadarshini, M. R. Panda, B. K. Mishra, Security in healthcare applications based on fog and cloud computing, Cyber Security in Parallel and Distributed Computing: Concepts,

Techniques, Applications and Case Studies (2019) 231–243.
[21] A. A. Abdellatif, A. Mohamed, C. F. Chiasserini, M. Tlili, A. Erbad, Edge computing for smart health: Context-aware approaches, opportunities, and challenges, IEEE Network 33 (2019) 196–203.
[22] A. Awad, A. Mohamed, C.-F. Chiasserini, T. Elfouly, Distributed in-network processing and resource optimization over mobile-health systems, Journal of Network and Computer Applications 82 (2017) 65–76.
[23] A. Pazienza, G. Mallardi, C. Fasciano, F. Vitulano, Artificial intelligence on edge computing: a healthcare scenario in ambient assisted living, in: Proceedings of the 5th Italian Workshop on Artificial Intelligence for Ambient Assisted Living 2019, co-located with 18th International Conference of the Italian Association for Artificial Intelligence, AI*AAL@AI*IA 2019, 2019, pp. 22–37.
[24] A. Pazienza, R. Anglani, G. Mallardi, C. Fasciano, P. Noviello, C. Tatulli, F. Vitulano, Adaptive critical care intervention in the internet of medical things, in: 2020 IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS), IEEE, 2020, pp. 1–8.
[25] P. Kumar, G. P. Gupta, R. Tripathi, An ensemble learning and fog-cloud architecture-driven cyber-attack detection framework for iomt networks, Computer Communications 166 (2021) 110–124.
[26] A. A. Hady, A. Ghubaish, T. Salman, D. Unal, R. Jain, Intrusion detection system for healthcare systems using medical and network data: A comparison study, IEEE Access 8 (2020) 106576–106584.
[27] M. A. Siddiqui, A. Fern, T. G. Dietterich, R. Wright, A. Theriault, D. W. Archer, Feedback-guided anomaly discovery via online optimization, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018, pp. 2200–2209.
[28] S. Das, W.-K. Wong, T. Dietterich, A. Fern, A. Emmott, Incorporating expert feedback into active anomaly discovery, in: 2016 IEEE 16th International Conference on Data Mining (ICDM), 2016, pp. 853–858.
[29] K. Veeramachaneni, I. Arnaldo, V. Korrapati, C. Bassias, K. Li, Aiˆ 2: training a big data machine to defend, in: 2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS), 2016, pp. 49–54.
[30] C. Ardito, T. Di Noia, E. Di Sciascio, D. Lofú, G. Mallardi, C. Pomo, F. Vitulano, Towards a trustworthy patient home-care thanks to an edge-node infrastructure, in: International Conference on Human-Centred Software Engineering, Springer, 2020, pp. 181–189.
[31] L. Viganò, D. Magazzeni, Explainable security, in: 2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), 2020, pp. 293–300.
[32] C. Ardito, T. D. Noia, C. Fasciano, D. Lofú, N. Macchiarulo, G. Mallardi, A. Pazienza, F. Vitulano, Towards a situation awareness for ehealth in ageing society, in: Proceedings of the Italian Workshop on Artificial Intelligence for an Ageing Society 2020 (AIxAS), co-located with 19th International Conference of the Italian Association for Artificial Intelligence (AIxIA 2020), 2020, pp. 40–55.
[33] Z. Chen, C. K. Yeo, B. S. Lee, C. T. Lau, Autoencoder-based network anomaly detection, in:

2018 Wireless Telecommunications Symposium (WTS), 2018, pp. 1–5.
[34] C. Zhou, R. C. Paffenroth, Anomaly detection with robust deep autoencoders, in: Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, 2017, pp. 665–674.
[35] M. Sakurada, T. Yairi, Anomaly detection using autoencoders with nonlinear dimensionality reduction, in: Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis, 2014, pp. 4–11.
[36] R. C. Aygun, A. G. Yavuz, Network anomaly detection with stochastically improved autoencoder based models, in: 2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud), 2017, pp. 193–198.
[37] I. Christov, G. Bortolan, Ranking of pattern recognition parameters for premature ventricular contractions classification by neural networks, Physiological Measurement 25 (2004) 1281.
[38] A. Hinneburg, D. A. Keim, M. Wawryniuk, Hd-eye: visual mining of high-dimensional data, IEEE Computer Graphics and Applications 19 (1999) 22–31.
[39] M. Kreuseler, T. Nocke, H. Schumann, A history mechanism for visual data mining, in: IEEE Symposium on Information Visualization, 2004, pp. 49–56.