pHealth 2024
M. Giacomini et al. (Eds.)

doi:10.3233/SHTI240068

Trustworthy Precision Medicine: An Interpretable Approach to Detecting Anomalous Behavior of IoT Devices

Gianni COSTA[a], Agostino FORESTIERO[a], Davide MACRÌ[a] and Riccardo ORTALE[a,1]
[a] *Institute for High-Performance Computing and Networking, National Research Council, Via P. Bucci 8-9C, Rende (CS), Italy*

Abstract. The growing integration of Internet of Things (IoT) technology within the healthcare sector has revolutionized healthcare delivery, enabling advanced personalized care and precise treatments. However, this raises significant challenges, demanding robust, intelligible, and effective monitoring mechanisms. We propose an interpretable machine- learning approach to the trustworthy and effective detection of behavioral anomalies within the realm of medical IoT. The discovered anomalies serve as indicators of potential system failures and security threats. Essentially, the detection of anomalies is accomplished by learning a classifier from the operational data generated by smart devices. The learning problem is dealt with in predictive association modeling, whose expressiveness and intelligibility enforce trustworthiness to offer a comprehensive, fully interpretable, and effective monitoring solution for the medical IoT ecosystem. Preliminary results show the effectiveness of our approach.

Keywords. Interpretable Artificial Intelligence, Precision Medicine, Medical IoT Security, Anomaly Detection

1. Introduction

The proliferation of the Internet of Medical Things (IoMT) has brought about revolutionary changes in healthcare, enabling remote monitoring, efficient data collection, and personalized patient care. However, integrating IoT devices in healthcare also poses significant challenges, necessitating robust monitoring strategies to safeguard sensitive medical information, guarantee the reliable operation of medical devices, and uphold patient confidentiality [8]. IoT-based precision medicine encompasses complex systems often consisting of smart devices with low processing capability, limited radio bandwidth, and battery resources, that provide volumes of data with rapid response times. Two primary challenges that need to be addressed are the smooth operation and security of these systems. Ensuring the smooth operation of such systems is of paramount importance. Any disruption can have significant consequences, making their seamless functioning a top priority. On the other hand, security threats are a persistent concern. Even though a security system may be effective at the time of deployment, it can become vulnerable over time as modern intruders and attackers continually evolve their strategies to avoid detection. Numerous intelligent approaches and platforms have been proposed

[1] Corresponding Author: Riccardo Ortale, riccardo.ortale@icar.cnr.it.

and developed to enhance information security in IoT environments. However, due to the unique characteristics of these systems, there is a need for innovative, advanced, and often custom-designed approaches to address security threats [13].

Anomaly detection [2], a critical field of data analysis, aims to identify un- usual or abnormal data, information, or behaviors. In distributed systems, such as IoMT, anomalies can be considered unusual/abnormal behaviors (or activities) of devices over time. Uncovering corresponding anomalous behavioral patterns is beneficial to revealing malfunctions, failure events, and unusual or anomalous actions of intruders or hacking tools [11].

In this paper, we propose an interpretable and effective approach to detecting anomalies in the operation of smart medical devices. The proposed approach enjoys the intelligibility of predictive association modeling, a seminal method from supervised machine learning, that integrates association-rule mining and classification [9]. Essentially, the anomalies in the behavior of medical IoT devices are dealt with from a classification perspective and carried out through rules uncovered using an Apriori-based search technique. This combines minimum and complement class support so that the number of parameters behind rule learning is limited to one, which is especially advantageous when classes are heavily skewed. The potentially very large set of uncovered rules is eventually distilled into a classifier through pruning. The latter is carried out as a covering procedure for overfitting avoidance, in which the F-measure is used to evaluate the predictive performance of rules over the training data. Thus, the only rules preserved are those enhancing the precision and recall of the resultant classifier across the classes. Preliminary comparison results against similar algorithms showed very interesting results.

2. The Proposed Approach

In this section, a method is presented for classifying device behaviors that relies on Class-Prediction Rule (CPR) to capture the associations between subsets of co-occurring devices and the discriminated classes. The reader is referred to [5] for a detailed formalization of notation, fundamental concepts, and technicalities. Due to space limitations, these details are omitted here. Model learning and prediction are the two stages of classification in our approach. The former constructs an associative classifier C from labeled device behaviors in a dataset. The latter utilizes C to make predictions about the behavior of unlabeled devices. Model learning, as detailed in Algorithm 1, takes four parameters: a dataset B of device behaviors, a set D of smart devices, a set L of various class labels in B, and a single global threshold τ for the identification of the minimum support thresholds relative to the different classes in L. FindCPRs is utilized to uncover a potentially-large collection **R** of CPRs devoted to separate the classes within L. Eventually, using the pruning procedure Distill, the set **R** of CPRs is reduced to a compact (and, hence, intelligible) associative classifier C.

FindCPRs implements an Apriori-based search for significant CPRs in the training data B. FindCPRs improves on the fundamental Apriori algorithm [1] through the incorporation of two effective mechanisms, namely *multiple minimum class support* [10] and *complement class support* [3], for distilling a suitable number of CPRs whose antecedents and consequents are positively correlated within each class in B. The aforesaid mechanisms are especially advantageous when the distribution of classes is

skewed in B. Indeed, without appropriate expedients, class imbalance generally obstacles the extraction of an appropriate number of CPRs from less frequently occurring classes, in addition to negatively acting on the correlation between the antecedents and consequents of CPRs, up to the point of yielding misleading (or, equivalently, negatively-correlated) CPRs.

FindCPRs may generate a huge number of CPRs, that likely overfit the training data and produce conflicting predictions since the associative patterns are inherently combinatorial. To circumvent these issues, an accurate classifier is extracted from **R** using Distill. It first sorts all available CPRs on the basis of the total order $\ll$, which is modeled after the one presented in [9]. Basically, given two CPRs $\mathbf{r}_i, \mathbf{r}_j \in \mathbf{R}$, $\mathbf{r}_i$ precedes $\mathbf{r}_j$, which is formalized as $\mathbf{r}_i \ll \mathbf{r}_j$, if (*i*) *conf* ($\mathbf{r}_i$) is higher than *conf* ($\mathbf{r}_j$), or (*ii*) *conf* ($\mathbf{r}_i$) equals *conf* (r_j), but *supp*($\mathbf{r}_i$) is higher than *supp*($\mathbf{r}_j$), or (*iii*) *conf* ($\mathbf{r}_i$) is identical to *conf* (r_j) and *supp*($\mathbf{r}_i$) equals *supp*($\mathbf{r}_j$), but *length*($\mathbf{r}_i$) is lower than *length*($\mathbf{r}_j$). Here, we refer to the length of a generic CPR $\mathbf{r} : \mathbf{D} \rightarrow c$ as the number of devices within the antecedent **r**, i.e., *length*(**r**) = |**D**|. In the case that $\mathbf{r}_i, \mathbf{r}_j$ share the same confidence, support, and length, $\mathbf{r}_i \ll \mathbf{r}_j$ still holds, provided that $\mathbf{r}_i$ was formed before $\mathbf{r}_j$.

Afterward, a covering procedure yields a compact classifier C composed of the fewest possible CPRs from **R** that achieve high predictive accuracy on unlabeled device behaviors. The covering procedure heuristically aims to maximize the effectiveness F(C) attained by the resulting C over all individual classes. F(C) is defined as [5]:

$$\mathrm{F(C)} = 1/|\mathrm{L}| * \sum_{\mathrm{c}\in\mathrm{L}} \mathrm{F}^{(\mathrm{c})}(\mathrm{C})$$

where $\mathrm{F}^{(\mathrm{c})}$(C) denotes the effectiveness (or, alternatively, predictive performance) of C relative to the generic class c, as defined below. defined below. F(C) gives the same weight to the effectiveness of C over all classes, in spite of their respective occurrence frequencies within the training data. This is particularly advantageous in the case of imbalanced classes, as F(C) is thus not ruled by the predictive performance of C across the most often occurring classes of devices behaviors. The covering procedure increments F(C) by acting separately on each $\mathrm{F}^{(c)}$(C), via the selection of CPRs from **R** that, once appended to C, enhance the predictive capability of the resulting classifier over class *c*. Assume C is the associative classifier shaped by the model learning phase, and U is a dataset of unlabeled device behaviors. C predicts the class C($\mathbf{b_t}$) for the arbitrary devices behavior $\mathbf{b}_t$ in U. C($\mathbf{b_t}$) is provided by the first CPR in C that covers $\mathbf{b_t}$, in order to avoid contradictory predictions from different triggered CPRs.

Algorithm 1 Model-Learning(B, D, L, τ)

Input: a training dataset B of devices behaviors;
a set D of smart devices;
a set L of class labels in B;
a support threshold τ;
Output: A classifier C= {$r_1 \vee \ldots \vee r_k$};
1: **R** ← ∅;
2: **B**′ ← ∅;
3: **for each** $b_t \in$ B **do**
4: b ← {d|d ∈ D, d is on over b_t};
5: **B**′ ← **B**′ ∪ {b};
6: **end for**
7: **R** ← FindCPRs(**B**′, D, L, τ);
8: C ← Distill(**R**, **B**′, L);
9: RETURN C

3. Preliminary Evaluation

We conducted a very preliminary evaluation comparing the performance of our approach in terms of precision, recall, and F-measure against two well-known rule-based baseline classifiers, One-R and LAC, using a publicly available IoT dataset, the smart-home Kyoto apartment collection [12]. OneR [7] is a rule-based classifier that builds a rule for each attribute in the training set before selecting the rule with the lowest error rate as its single rule, while LAC - lazy associative classifier - [14] produces classification association rules specific to each test tuple, unlike traditional associative classifiers, which focus on a collection of ranked classification association rules from the training data. As shown in Figure 1, our strategy outperforms all other approaches on the considered dataset.

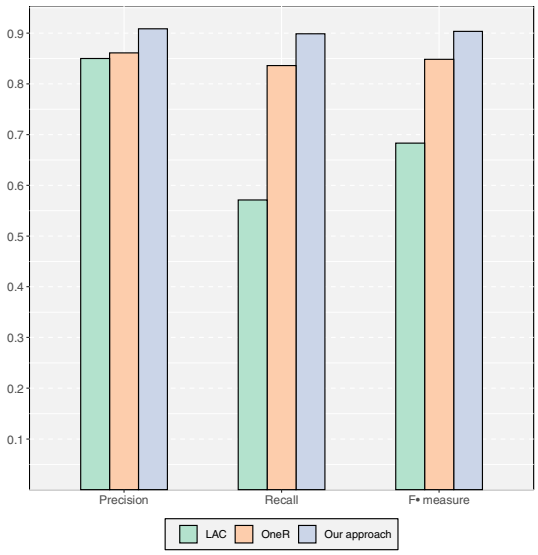

Figure 1. Comparison on Kyoto dataset

4. Conclusion

We proposed a machine-learning approach to the trustworthy detection of IoT behavioral anomalies that involves the classification of traces of smart device operation through predictive association modeling. Its peculiar design allows IoT threat detection to be governed by only one parameter. The results of preliminary comparative tests on real-world benchmark data showed the effectiveness of the proposed approach. Our proposal aims to emphasize the importance of user trust and comprehension in the deployment of monitoring solutions, thereby marking a significant step forward in the development of robust, interpretable, and user-friendly medical IoT systems. Exploring the utilization of co-clustering techniques [6] for smart devices could yield an understanding of relationships among devices exhibiting similar behavior. Such insights would be valuable for further behavioral analysis, more accurately recommending [4] and enforcing specific security policies,

and optimizing resource allocation, which are of significant practical importance in precision medicine.

References

[1] R. Agrawal and R. Srikant. Fast algorithms for mining association rules. In *Proc. of Int. Conf. on Very Large Data Bases*, pages 487 – 499, 1994.
[2] Mohiuddin Ahmed, Abdun Naser Mahmood, and Jiankun Hu. A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*, 60:19–31, 2016.
[3] B. Arunasalam and S. Chawla. CCCS: A top-down association classifier for imbalanced class distribution. In *Proceedings of ACM SIGKDD International Conference on Knowl- edge Discovery and Data Mining*, pages 517–522, 2006.
[4] N. Barbieri, G. Costa, G. Manco, and R. Ortale. Modeling item selection and relevance for accurate recommendations: a bayesian approach. In *Proceedings of the ACM Conference on Recommender Systems*, pages 21 – 28, 2011.
[5] G. Costa, A. Forestiero, and R. Ortale. Rule-based detection of anomalous patterns in device behavior for explainable iot security. *IEEE Transactions on Services Computing*, 16(6):4514 – 4525, 2023.
[6] G. Costa, G. Manco, and R. Ortale. A hierarchical model-based approach to co-clustering high-dimensional data. In *Proceedings of the ACM Symposium on Applied Computing*, pages 886 – 890, 2008.
[7] R.C. Holte. Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, 11:63–91, 1993.
[8] Nan Li, Minxian Xu, Qimeng Li, Jikui Liu, Shudi Bao, Ye Li, Jianzhong Li, and Hairong Zheng. A review on security issues and solutions for precision health in internet-of-medical- things systems. *Security and Safety*, 2022.
[9] B. Liu, W. Hsu, and Y. Ma. Integrating classification and association rule mining. In *Proceedings of ACM SIGKDD International Conference on Kwnoledge Discovery and Data Mining*, pages 80–86, 1998.
[10] B. Liu, Y. Ma, and C.K. Wong. Improving an association rule based classifier. In *Pro- ceedings of Principles of Data Mining and Knowledge Discovery*, pages 504–509, 2000.
[11] In˜es Martins, Jo˜ao S. Resende, Patr´ıcia R. Sousa, Sim˜ao Silva, Lu´ıs Antunes, and Jo˜ao Gama. Host-based ids: A review and open issues of an anomaly detection system in iot. *Future Generation Computer Systems*, 133:95–113, 2022.
[12] S. Szewcyzk, K. Dwan, B. Minor, B. Swedlove, and D. Cook. Annotating smart environ- ment sensor data for activity learning. *Technology and health care : official journal of the European Society for Engineering and Medicine*, 17(3):161–169, 2009.
[13] Chandra Thapa and Seyit Camtepe. Precision health data: Requirements, challenges and existing techniques for data security and privacy. *Computers in biology and medicine*, 2021.
[14] Adriano Veloso, Wagner Meira, and Mohammed J. Zaki. Lazy associative classification. In *Sixth International Conference on Data Mining (ICDM'06)*, pages 645–654, 2006.