Differentially Private Federated Learning for Anomaly Detection in eHealth Networks

Ana Cholakoska*
Ss. Cyril and Methodius University
Faculty of Electrical Engineering and Information Technologies
Skopje, North Macedonia
acholak@feit.ukim.edu.mk

Bjarne Pfitzner*
Hasso Plattner Institute
Digital Health — Connected Healthcare
Potsdam, Germany
bjarne.pfitzner@hpi.de

Hristijan Gjoreski
Ss. Cyril and Methodius University
Faculty of Electrical Engineering and Information Technologies
Skopje, North Macedonia
hristijang@feit.ukim.edu.mk

Valentin Rakovic
Ss. Cyril and Methodius University
Faculty of Electrical Engineering and Information Technologies
Skopje, North Macedonia
valentin@feit.ukim.edu.mk

Bert Arnrich
Hasso Plattner Institute
Digital Health — Connected Healthcare
Potsdam, Germany
bert.arnrich@hpi.de

Marija Kalendar
Ss. Cyril and Methodius University
Faculty of Electrical Engineering and Information Technologies
Skopje, North Macedonia
marijaka@feit.ukim.edu.mk

ABSTRACT

Increasing number of ubiquitous devices are being used in the medical field to collect patient information. Those connected sensors can potentially be exploited by third parties who want to misuse personal information and compromise the security, which could ultimately result even in patient death. This paper addresses the security concerns in eHealth networks and suggests a new approach to dealing with anomalies. In particular we propose a concept for safe in-hospital learning from internet of health things (IoHT) device data while securing the network traffic with a collaboratively trained anomaly detection system using federated learning. That way, real time traffic anomaly detection is achieved, while maintaining collaboration between hospitals and keeping local data secure and private. Since not only the network metadata, but also the actual medical data is relevant to anomaly detection, we propose to use differential privacy (DP) for providing formal guarantees of the privacy spending accumulated during the federated learning.

CCS CONCEPTS

• **Security and privacy** → **Intrusion detection systems**; • **Applied computing** → *Health care information systems*; • **Computing methodologies** → *Anomaly detection*; **Distributed artificial intelligence**.

KEYWORDS

anomaly detection, federated learning, eHealth

*Both authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
UbiComp '21, September 21–26, 2021, All Over the World
© 2021 Association for Computing Machinery.
https://doi.org/10.1145/1122445.1122456

ACM Reference Format:
Ana Cholakoska, Bjarne Pfitzner, Hristijan Gjoreski, Valentin Rakovic, Bert Arnrich, and Marija Kalendar. 2021. Differentially Private Federated Learning for Anomaly Detection in eHealth Networks. In *Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

The rapid development of the internet of things (IoT) leads to a variety of applications in society, enabling simplification and improvement of the quality of life of end-users. This is also the case with mobile and eHealth, where human health and well-being come first. With the help of various sensors for monitoring the human body, the possibilities for diagnosis, early prevention, treatment and administration of drugs are becoming faster and easier [15]. Also, mobile devices and smartwatches with accelerometers and pulse oximeters today play a key role in the remote monitoring of patients and health evaluation of regular people (fall detection, etc.) [8].

However, such devices also have their drawbacks. Due to the heterogeneity of sensors and technologies used in the transmission of data in eHealth environments, the risk of violating the privacy of patients' data and their electronic health records increases. Networked devices can be turned off, reconfigured or reprogrammed, which could put patients at risk or have catastrophic consequences on their health [22]. For instance, it was shown that malware could be deployed to pacemakers or insulin pumps [3, 18] that could quickly result in the patient's death.

In order to preserve the safety of patients and their data, new and improved ways are being sought to detect such anomalies in real-time so that timely responses can be made. Because we are considering protecting IoT networks, the traditional network intrusion detection system (NIDS) that exist cannot fully cope with the new attacks that are taking place. Machine learning (ML) has already shown high efficacy in detecting anomalies. Recently, federated learning is emerging as a promising new variant that can significantly improve the time to detect and deal with such anomalies without compromising patient data. Hospitals can keep their

data on-site and only have to commit to training the model locally and sending model updates to a server. The collaboration of many parties allows the model to have access to a larger and more diverse dataset, making the predictions more accurate than classical, centralized ML.

This paper proposes securing the in-hospital network traffic with an anomaly detection system trained collaboratively using federated learning. The system takes network information, as well as patient data as input and uses differential privacy (DP) to ensure the data privacy during the training process. With the anomaly detection system trained an in place, the locally collected internet of health things (IoHT) data can be trusted and used for e.g. ML training.

2 BACKGROUND

2.1 IoT for eHealth

To meet the needs of patients, doctors and hospitals, IoHT devices must meet certain security requirements. However, the traditional network requirements like confidentiality, integrity and availability are not enough for these specialized types of devices. For IoHT systems to be considered safe, they must meet additional security requirements [17], such as:

- Privacy and proper use of data: all data collected and processed must be used according to rules in accordance to the GDPR law.
- Access control: only authorized users (medical personnel) should be allowed to have access to IoHT devices, as well as the possibility to modify some of their parameters.
- Anonymity: systems should be able to protect patients' privacy and their data.
- Authenticity: systems should have the ability to verify and validate user profiles.
- Data integrity: systems should be able to prevent unauthorized data modifications.

The possible attacks which may occur in such networks are explained in the following paragraphs.

Data confidentiality attacks. In these types of attacks, personal and private information like medical patient records is leaked, modified or hijacked. Here, different passive types of attacks can be performed: eavesdropping, wiretapping, packet capturing and data interception.

Privacy attacks. These types of attacks violate the location, behavior or real identities of patients. Using traffic analysis, identity tracking or location tracking, attackers can relate a person with a place, putting their privacy and maybe life at risk.

Message authentication and data integrity attacks. These attacks alter messages transmitted through the network in order to target the integrity of a system or data. By cloning, spoofing, malicious script injection, message tampering and alteration, or malicious data injection, attackers can manipulate sent/received messages or send false messages to doctors, which can lead to accidents.

Device and user authentication attacks. Attackers use replay attacks, masquerading, cracking, dictionary, rainbow table, birthday, session hijacking, brute force and man in the middle attacks to overcome passwords and gain access to patients' credentials and hospital sensitive information for fraud and other purposes.

Malware attacks. These types of attacks exploit a software vulnerability or security gap to gain unauthorized access to the medical system, delete or modify patient sensitive information. Examples are spyware, ransomware, botnet attacks, logic bombs, remote access Trojans and worms.

Availability attacks. By reducing the performance of medical devices and systems, attackers can prevent nurses and doctors from gaining a real-time insight in a patient's condition, which can degrade their health and possibly lead to false diagnosis and incorrect therapy. Also, they can prevent devices from being operational — temporarily or permanently, exhausting the system's resources. Here, different types of attacks can be performed: denial of service, deauthentication, wireless jamming, flooding (ICMP flooding, SYN flooding), black nurse and delay.

2.2 Federated Learning

Federated learning is a novel distributed ML approach enabling the training of models on private datasets, i.e., datasets that should not be transferred and shared due to privacy reasons. The collaboration between the participants is enabled by exchanging model parameters instead of the actual sensitive data itself. [9]

The training procedure involves a central server component which keeps track of the current global model, typically an artificial neural network, and controls the training. In each federated training round, the server selects a subset of clients at random who receive the current model parameters. This model is then trained for a predefined number of local epochs by applying a gradient-based optimization method using the local datasets before sending the updated model parameters back to the server. The global model at the server is finally updated using the average of all received parameters, weighted by the amount of local data. This process is repeated until model convergence is achieved, or alternatively in an online learning scenario the training is performed indefinitely. [9]

2.2.1 *Federated Learning with Differential Privacy.*

DP describes a formal formulation of privacy constraints in data science and ML [4]. It provides upper bounds for the risk of re-identifying the impact of a single datum in database queries or ML models. Formally, DP is defined as follows.

Definition 2.1 ((ϵ, δ)-DP [4]). A randomized mechanism $\mathcal{M} : \mathcal{D} \rightarrow \mathcal{R}$ with domain $\mathcal{D}$ and range $\mathcal{R}$ satisfies (ϵ, δ)-DP if for any two adjacent inputs $d, d' \in \mathcal{D}$ and for any subset of outputs $\mathcal{S} \subseteq \mathcal{R}$ it holds that,

$$\Pr[\mathcal{M}(d) \in S] \leq e^{\epsilon} \Pr[\mathcal{M}(d') \in S] + \delta$$

Adjacent inputs are defined as inputs differing in a single instance, that is some data point is present in one and absent in the other. The privacy loss, ϵ, is the upper bound of the distance of two mechanism outputs generated by adjacent inputs. It is also the logarithm of the ratio of probabilities of observing the same output with two adjacent inputs. For (ϵ, δ)-DP, also called approximate DP, δ represents the probability that the aforementioned bound does not hold. Applied to ML, DP entails the introduction of noise into the training procedure. Typically, the gradients are perturbed with Gaussian noise having zero mean and variance tuned to the sensitivity of the gradients. Multiple training rounds increase the privacy loss and Abadi et al. [1] propose the use of a *moments accountant* to

keep track of the current privacy spending. Recently, this approach has been extended to federated learning, hiding the participation of individual clients or single data points in the training process [7, 21]. It has shown to provide protection against membership inference and data reconstruction attacks [7, 10].

3 RELATED WORK

3.1 Anomaly Detection in eHealth Networks

One of the most common approaches when it comes to IoHT traffic anomaly detection is using NIDSs or host-based intrusion detection systems (IDSs) (if the logs and data of the sensors are also being used). Both approaches use different metrics to separate *signatures* (known attacks) and *anomalies* (unknown attacks) from normal network traffic flow. The signature-based approaches can easily detect known patterns, but lack the ability of detecting new attacks [11]. By learning from benign network data, the anomaly-based approaches perform better because they are able to detect unknown attacks as well as known ones. However, this comes with greater computational cost and less detection accuracy.

The recent advancements of ML techniques have proven beneficial in addressing IoHT anomaly detection. Some of the approaches to secure IoHT with ML involve sensor anomaly detection, while other approaches used ML for intrusion and malware detection.

Gao and Thamilarasu [6] used decision trees, support vector machines (SVMs) and K-means clustering to detect attacks on implantable devices. The results showed that the decision tree algorithm achieved the highest accuracy with fast training and prediction compared to the other algorithms. The SVM was the algorithm of choice in the study of Verner and Butvinik [20], where data of a blood glucose sensor was inspected in an attempt to detect accidental data modification intrusions. Other researchers [13] implemented an ML model to separate valid from anomalous data using a combination of a neural network (NN) with Ensemble Linear Regression as detection method.

Shakeel P et al. [19] used a Deep-Q-Network methodology, where the IoHT system was analyzed by a deep NN in order to detect and eliminate any malware attacks. Malware detection-based research can be found in the work of Fernández Maimó et al. [5], where Naive Bayes and one-class-SVM techniques were used to detect and classify ransomware. Alrashdi et al. [2] were able to obtain better accuracy and detection time of their decentralized fog-attack detection architecture compared to a centralized framework using an online sequential extreme learning machine.

3.2 Anomaly Detection using Federated Learning

Recently, federated learning has been applied to anomaly detection in order to benefit from larger databases and faster response times as opposed to anomaly detection hosted as a service.

Preuveneers et al. [16] propose a combination of federated learning and distributed ledger technology. They train an autoencoder model to detect anomalies in network traffic. The use of a blockchain removes the need for a trusted central entity and subsequently the risks of a single point of failure.

The authors in [14] developed a device-type-specific IDS that converts data into language symbols and uses language analysis for anomaly detection. In addition, the security gateways collaborate in a federated learning system and do not require data labeling. Their evaluation showed a detection rate of 95.6% and no false positives.

Zhao et al. [23] extend the anomaly detection task with two more (VPN or Tor traffic recognition and traffic classification) and solve all problems simultaneously utilizing a multi-task learning approach. Their multi-task deep neural network is able to outperform centralized models on all tasks.

As a means to further improve the detection accuracy, [12] propose a random forest ensemble of multiple Gated Recurrent Units trained with different window sizes. This creates a trade-off between the overhead of training multiple models in parallel and the performance improvement of the ensemble approach.

Current related work does not specifically target the medical or IoHT domain. As data here is very sensitive, it is vital to consider data privacy when developing solutions for anomaly detection in this area. Federated learning is the first step towards that goal, but there is room for improvement.

4 PROPOSED SYSTEM ARCHITECTURE

We propose to enable safe in-hospital learning from IoHT device data by securing the network traffic with an anomaly detection system trained collaboratively using federated learning. This approach will enable real time traffic anomaly detection, while maintaining local patient data intact and implementing collaboration between hospitals. Fig. 1 shows the proposed system architecture. It consists of multiple clients in the form of hospitals or clinics, and a server that can communicate with all clients.

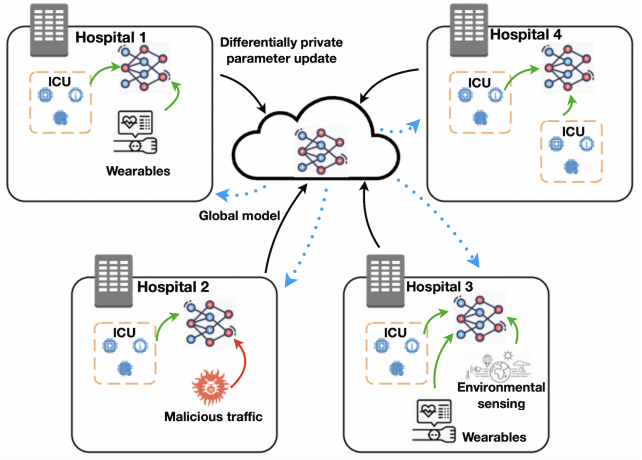

Figure 1: Proposed system for eHealth anomaly detection with federated learning. Different hospitals possess different kinds of data sources. All network traffic is used to train the anomaly detection model, and the model is used to detect malicious traffic.

Client-side. Hospitals collect data with a number of connected IoHT devices. Intensive care units (ICUs) are equipped with stationary patient monitoring devices measuring heart rate, blood pressure, blood oxygen saturation, etc., but also active devices such as infusion pumps and ventilators. Additionally, more and more

wearable devices are introduced into the normal wards in order to collect more patient data e.g. during surgery recovery. Furthermore, clinics may possess environment monitoring sensors such as humidity, temperature or smoke sensors. All that data can provide feedback to the hospital staff and possibly alert them in case of any signs of condition deterioration, which makes it important that the information is accurate and not compromised. Moreover, the hospitals may have an interest in training ML models with this data, which also requires a clean and high-quality dataset. As a means to protect the data integrity and network systems, we propose that hospitals collaborate in a federated learning system to jointly train an anomaly detection model. Different models could be applied here, e.g. autoencoders or k-nearest-neighbors. When notified by the parameter server, the hospital has to update the model with its local data and send the updates back to the server. Including the actual data in the input to the model is beneficial, since there could be an attempt of malicious data injection. Since the data could be sensitive patient data, it has to be protected from membership inference or data reconstruction attacks on federated learning systems. Here, the introduction of DP provides privacy guarantees to data owners. It is to be investigated if the trade-off between privacy and model accuracy is feasible for the use-case.

Server-side. The server takes on the role of the parameter server in the federated learning system. The global model will be initiated here, and the server deals with hospital selection per training round, as well as parameter distribution, collection and averaging.

The proposed federated learning architecture has several advantages for the participating hospitals. They can benefit from each others benign train data, making the anomaly detection model training set larger, and the model itself stronger. This can remove reservations for introducing more ubiquitous devices into everyday clinical practice, which can benefit the patients as well as doctors or nurses. The former receive better care while the latter have more information at their disposal to determine appropriate interventions and treatments.

5 CONCLUSION AND FUTURE WORK

This paper describes an anomaly detection system for eHealth, enabling the secure use of network traffic generated by IoHT devices, as well as the patient data they generate. Federated learning with DP is the method of choice to jointly train the anomaly detection model and protect the sensitive patient information in the process.

Future work will focus on implementing the proposed system and evaluating the trade-off between privacy and accuracy imposed by DP. Moreover, future work will strive to address issues related to the non-independent and identical data distribution across hospitals and patients. Naturally occurring outliers in patient data, e.g. when their condition is deteriorating, have to be distinguished from malicious data injection.

ACKNOWLEDGMENTS

This work has been supported by the WideHealth project - European Union's Horizon 2020 research and innovation programmender grant agreement No. 95227.

This research was partly funded by the Federal Ministry of Education and Research of Germany in the framework of KI-LAB-ITSE (project number 01IS19066).

REFERENCES

[1] Martin Abadi, Andy Chu, Ian Goodfellow, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, CCS '16, page 308–318, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450341394. doi: 10.1145/2976749.2978318. URL https://doi.org/10.1145/2976749.2978318.

[2] Ibrahim Alrashdi, Ali Alqazzaz, Raed Alharthi, Esam Aloufi, Mohamed A. Zohdy, and Hua Ming. Fbad: Fog-based attack detection for iot healthcare in smart cities. In *2019 IEEE 10th Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, pages 0515–0522, 2019. doi: 10.1109/UEMCON47517.2019.8992963.

[3] Jake L Beavers, Michael Faulks, and Jims Marchang. Hacking nhs pacemakers: A feasibility study. In *2019 IEEE 12th International Conference on Global Security, Safety and Sustainability (ICGS3)*, pages 206–212, 2019. doi: 10.1109/ICGS3.2019.8688214.

[4] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3–4):211–407, August 2014. ISSN 1551-305X. doi: 10.1561/0400000042. URL https://doi.org/10.1561/0400000042.

[5] Lorenzo Fernández Maimó, Alberto Huertas Celdrán, Ángel L. Perales Gómez, Félix J. García Clemente, James Weimer, and Insup Lee. Intelligent and dynamic ransomware spread detection and mitigation in integrated clinical environments. *Sensors*, 19(5), 2019. ISSN 1424-8220. doi: 10.3390/s19051114. URL https://www.mdpi.com/1424-8220/19/5/1114.

[6] Sida Gao and Geethapriya Thamilarasu. Machine-learning classifiers for security in connected medical devices. In *2017 26th International Conference on Computer Communication and Networks (ICCCN)*, pages 1–5, 2017. doi: 10.1109/ICCCN.2017.8038507.

[7] Robin C. Geyer, Tassilo Klein, and Moin Nabi. Differentially private federated learning: A client level perspective. *CoRR*, abs/1712.07557, 2017. URL http://arxiv.org/abs/1712.07557.

[8] Hristijan Gjoreski, Mitja Lustrek, and Matjaz Gams. Accelerometer placement for posture recognition and fall detection. In *2011 Seventh International Conference on Intelligent Environments*, pages 47–54, 2011. doi: 10.1109/IE.2011.11.

[9] H. Brendan McMahan, Eider Moore, Daniel Ramage, and Blaise Agüera y Arcas. Federated learning of deep networks using model averaging. *CoRR*, abs/1602.05629, 2016. URL http://arxiv.org/abs/1602.05629.

[10] H. Brendan McMahan, Daniel Ramage, Kunal Talwar, and Li Zhang. Learning differentially private language models without losing accuracy. *CoRR*, abs/1710.06963, 2017. URL http://arxiv.org/abs/1710.06963.

[11] Akhil Jabbar Meera, M. V. V. Prasad Kantipudi, and Rajanikanth Aluvalu. Intrusion detection system for the iot: A comprehensive review. In Ajith Abraham, M. A. Jabbar, Sanju Tiwari, and Isabel M. S. Jesus, editors, *Proceedings of the 11th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2019)*, pages 235–243, Cham, 2021. Springer International Publishing. ISBN 978-3-030-49345-5.

[12] Viraaji Mothukuri, Prachi Khare, Reza M. Parizi, Seyedamin Pouriyeh, Ali Dehghantanha, and Gautam Srivastava. Federated learning-based anomaly detection for iot security attacks. *IEEE Internet of Things Journal*, pages 1–1, 2021. doi: 10.1109/JIOT.2021.3077803.

[13] Sumit Kumar Nagdeo and Judhistir Mahapatro. Wireless body area network sensor faults and anomalous data detection and classification using machine learning. In *2019 IEEE Bombay Section Signature Conference (IBSSC)*, pages 1–6, 2019. doi: 10.1109/IBSSC47189.2019.8973004.

[14] Thien Duc Nguyen, Samuel Marchal, Markus Miettinen, Minh Hoang Dang, N. Asokan, and Ahmad-Reza Sadeghi. Dïot: A crowdsourced self-learning approach for detecting compromised iot devices. *CoRR*, abs/1804.07474, 2018. URL http://arxiv.org/abs/1804.07474.

[15] C. V. Nisha Angeline, S. Muthuramlingam, E. Rahul Ganesh, S. Siva Pratheep, and V. Nishanthan. Medical iot—automatic medical dispensing machine. In E. S. Gopi, editor, *Machine Learning, Deep Learning and Computational Intelligence for Wireless Communication*, pages 323–330, Singapore, 2021. Springer Singapore. ISBN 978-981-16-0289-4.

[16] Davy Preuveneers, Vera Rimmer, Ilias Tsingenopoulos, Jan Spooren, Wouter Joosen, and Elisabeth Ilie-Zudor. Chained anomaly detection models for federated learning: An intrusion detection case study. *Applied Sciences*, 8(12), 2018. ISSN 2076-3417. doi: 10.3390/app8122663. URL https://www.mdpi.com/2076-3417/8/12/2663.

[17] Panagiotis I. Radoglou Grammatikis, Panagiotis G. Sarigiannidis, and Ioannis D. Moscholios. Securing the internet of things: Challenges, threats and solutions. *Internet of Things*, 5:41–70, 2019. ISSN 2542-6605. doi: https://doi.org/10.1016/j.iot.2018.11.003. URL https://www.sciencedirect.com/science/

article/pii/S2542660518301161.

[18] Muhammad Muneeb Ur Rehman, Hafiz Zia Ur Rehman, and Zeashan Hameed Khan. Cyber-attacks on medical implants: A case study of cardiac pacemaker vulnerability. *International Journal of Computing and Digital Systems*, 9(6):1229–1235, 2020.

[19] Mohamed Shakeel P, Baskar .S, V.R.Sarma Dhulipala, Sukumar Mishra, and Mustafa Jaber. Maintaining security and privacy in health care system using learning based deep-q-networks. *Journal of Medical Systems*, 42, 08 2018. doi: 10.1007/s10916-018-1045-z.

[20] Alexander Verner and Dany Butvinik. A machine learning approach to detecting sensor data modification intrusions in wbans. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 161–169, 2017. doi: 10.1109/ICMLA.2017.0-163.

[21] Kang Wei, Jun Li, Ming Ding, Chuan Ma, Howard H. Yang, Farhad Farokhi, Shi Jin, Tony Q. S. Quek, and H. Vincent Poor. Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security*, 15:3454–3469, 2020. doi: 10.1109/TIFS.2020.2988575.

[22] Jean-Paul A. Yaacoub, Mohamad Noura, Hassan N. Noura, Ola Salman, Elias Yaacoub, Raphaël Couturier, and Ali Chehab. Securing internet of medical things systems: Limitations, issues and recommendations. *Future Generation Computer Systems*, 105:581–606, 2020. ISSN 0167-739X. doi: https://doi.org/10.1016/j.future.2019.12.028. URL https://www.sciencedirect.com/science/article/pii/S0167739X1930568.

[23] Ying Zhao, Junjun Chen, Di Wu, Jian Teng, and Shui Yu. Multi-task network anomaly detection using federated learning. In *Proceedings of the Tenth International Symposium on Information and Communication Technology*, SoICT 2019, page 273–279, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450372459. doi: 10.1145/3368926.3369705. URL https://doi.org/10.1145/3368926.3369705.

Received June 2021; accepted July 2021