http://www.icact.org/journal.asp

ICACT-TACT JOURNAL

Transactions on Advanced Communications Technology

Volume 7 Issue 5, September. 2018, ISSN: 2288-0003

Editor-in-Chief
Prof. Thomas Byeongnam YOON, PhD.

GIRI
Global IT Research Institute

Journal Editorial Board

■ Editor-in-Chief

Prof. Thomas Byeongnam YOON, PhD.
Founding Editor-in-Chief
ICACT Transactions on the Advanced Communications Technology (TACT)

■ Editors

Prof. Jun-Chul Chun, Kyonggi University, Korea
Dr. JongWon Kim, GIST (Gwangju Institute of Science & Technology), Korea
Dr. Xi Chen, State Grid Corparation of China, China
Prof. Arash Dana, Islamic Azad university , Central Tehran Branch, Iran
Dr. Pasquale Pace, University of Calabria - DEIS - Italy, Italy
Dr. Mitch Haspel, Stochastikos Solutions R&D, Israel
Prof. Shintaro Uno, Aichi University of Technology, Japan
Dr. Tony Tsang, Hong Kong Polytechnic UNiversity, Hong Kong
Prof. Kwang-Hoon Kim, Kyonggi University, Korea
Prof. Rosilah Hassan, Universiti Kebangsaan Malaysia(UKM), Malaysia
Dr. Sung Moon Shin, ETRI, Korea
Dr. Takahiro Matsumoto, Yamaguchi University, Japan
Dr. Christian Esteve Rothenberg, CPqD - R&D Center for. Telecommunications, Brazil
Prof. Lakshmi Prasad Saikia, Assam down town University, India
Prof. Moo Wan Kim, Tokyo University of Information Sciences, Japan
Prof. Yong-Hee Jeon, Catholic Univ. of Daegu, Korea
Dr. E.A.Mary Anita, Prathyusha Institute of Technology and Management, India
Dr. Chun-Hsin Wang, Chung Hua University, Taiwan
Prof. Wilaiporn Lee, King Mongkut's University of Technology North, Thailand
Dr. Zhi-Qiang Yao, XiangTan University, China
Prof. Bin Shen, Chongqing Univ. of Posts and Telecommunications (CQUPT), China
Prof. Vishal Bharti, Dronacharya College of Engineering, India
Dr. Marsono, Muhammad Nadzir , Universiti Teknologi Malaysia, Malaysia
Mr. Muhammad Yasir Malik, Samsung Electronics, Korea
Prof. Yeonseung Ryu, Myongji University, Korea
Dr. Kyuchang Kang, ETRI, Korea
Prof. Plamena Zlateva, BAS(Bulgarian Academy of Sciences), Bulgaria
Dr. Pasi Ojala, University of Oulu, Finland
Prof. CheonShik Kim, Sejong University, Korea
Dr. Anna bruno, University of Salento, Italy
Prof. Jesuk Ko, Gwangju University, Korea
Dr. Saba Mahmood, Air University Islamabad Pakistan, Pakistan
Prof. Zhiming Cai, Macao University of Science and Technology, Macau
Prof. Man Soo Han, Mokpo National Univ., Korea
Mr. Jose Gutierrez, Aalborg University, Denmark

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Dr. Youssef SAID, Tunisie Telecom, Tunisia
Dr. Noor Zaman, King Faisal University, Al Ahsa Hofuf, Saudi Arabia
Dr. Srinivas Mantha, SASTRA University, Thanjavur, India
Dr. Shahriar Mohammadi, KNTU University, Iran
Prof. Beonsku An, Hongik University, korea
Dr. Guanbo Zheng, University of Houston, USA
Prof. Sangho Choe, The Catholic University of Korea, korea
Dr. Gyanendra Prasad Joshi, Yeungnam University, korea
Dr. Tae-Gyu Lee, Korea Institue of Industrial Technology(KITECH), korea
Prof. Ilkyeun Ra, University of Colorado Denver, USA
Dr. Yong Sun, Beijing University of Posts and Telecommunications, China
Dr. Yulei Wu, Chinese Academy of Sciences, China
Mr. Anup Thapa, Chosun University, korea
Dr. Vo Nguyen Quoc Bao, Posts and Telecommunications Institute of Technology, Vietnam
Dr. Harish Kumar, Bhagwant institute of technology, India
Dr. Jin REN, North china university of technology, China
Dr. Joseph Kandath, Electronics & Commn Engg, India
Dr. Mohamed M. A. Moustafa, Arab Information Union (AIU), Egypt
Dr. Mostafa Zaman Chowdhury, Kookmin University, Korea
Prof. Francis C.M. Lau, Hong Kong Polytechnic University, Hong Kong
Prof. Ju Bin Song, Kyung Hee University, korea
Prof. KyungHi Chang, Inha University, Korea
Prof. Sherif Welsen Shaker, Kuang-Chi Institute of Advanced Technology, China
Prof. Seung-Hoon Hwang, Dongguk University, Korea
Prof. Dal-Hwan Yoon, Semyung University, korea
Prof. Chongyang ZHANG, Shanghai Jiao Tong University, China
Dr. H K Lau, The Open University of Hong Kong, Honh Kong
Prof. Ying-Ren Chien, Department of Electrical Engineering, National Ilan University, Taiwan
Prof. Mai Yi-Ting, Hsiuping University of Science and Technology, Taiwan
Dr. Sang-Hwan Ryu, Korea Railroad Research Institute, Korea
Dr. Yung-Chien Shih, MediaTek Inc., Taiwan
Dr. Kuan Hoong Poo, Multimedia University, Malaysia
Dr. Michael Leung, CEng MIET SMIEEE, Hong Kong
Dr. Abu sahman Bin mohd Supa'at, Universiti Teknologi Malaysia, Malaysia
Prof. Amit Kumar Garg, Deenbandhu Chhotu Ram University of Science & Technology, India
Dr. Jens Myrup Pedersen, Aalborg University, Denmark
Dr. Augustine Ikechi Ukaegbu, KAIST, Korea
Dr. Jamshid Sangirov, KAIST, Korea
Prof. Ahmed Dooguy KORA, Ecole Sup. Multinationale des Telecommunications, Senegal
Dr. Se-Jin Oh, Korea Astronomy & Space Science Institute, Korea
Dr. Rajendra Prasad Mahajan, RGPV Bhopal, India
Dr. Woo-Jin Byun, ETRI, Korea
Dr. Mohammed M. Kadhum, School of Computing, Goodwin Hall, Queen's University , Canada
Prof. Seong Gon Choi, Chungbuk National University, Korea
Prof. Yao-Chung Chang, National Taitung University, Taiwan
Dr. Abdallah Handoura, Engineering school of Gabes - Tunisia, Tunisia
Dr. Gopal Chandra Manna, BSNL, India

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Dr. Il Kwon Cho, National Information Society Agency, Korea
Prof. Jiann-Liang Chen, National Taiwan University of Science and Technology, Taiwan
Prof. Ruay-Shiung Chang, National Dong Hwa University, Taiwan
Dr. Vasaka Visoottiviseth, Mahidol University, Thailand
Prof. Dae-Ki Kang, Dongseo University, Korea
Dr. Yong-Sik Choi, Research Institute, IDLE co., ltd, Korea
Dr. Xuena Peng, Northeastern University, China
Dr. Ming-Shen Jian, National Formosa University, Taiwan
Dr. Soobin Lee, KAIST Institute for IT Convergence, Korea
Prof. Yongpan Liu, Tsinghua University, China
Prof. Chih-Lin HU, National Central University, Taiwan
Prof. Chen-Shie Ho, Oriental Institute of Technology, Taiwan
Dr. Hyoung-Jun Kim, ETRI, Korea
Prof. Bernard Cousin, IRISA/Universite de Rennes 1, France
Prof. Eun-young Lee, Dongduk Woman s University, Korea
Dr. Porkumaran K, NGP institute of technology India, India
Dr. Feng CHENG, Hasso Plattner Institute at University of Potsdam, Germany
Prof. El-Sayed M. El-Alfy, King Fahd University of Petroleum and Minerals, Saudi Arabia
Prof. Lin You, Hangzhou Dianzi Univ, China
Mr. Nicolai Kuntze, Fraunhofer Institute for Secure Information Technology, Germany
Dr. Min-Hong Yun, ETRI, Korea
Dr. Seong Joon Lee, Korea Electrotechnology Research Institute, korea
Dr. Kwihoon Kim, ETRI, Korea
Dr. Jin Woo HONG, Electronics and Telecommunications Research Inst., Korea
Dr. Heeseok Choi, KISTI(Korea Institute of Science and Technology Information), korea
Dr. Somkiat Kitjongthawonkul, Australian Catholic University, St Patrick's Campus, Australia
Dr. Dae Won Kim, ETRI, Korea
Dr. Ho-Jin CHOI, KAIST(Univ), Korea
Dr. Su-Cheng HAW, Multimedia University, Faculty of Information Technology, Malaysia
Dr. Myoung-Jin Kim, Soongsil University, Korea
Dr. Gyu Myoung Lee, Institut Mines-Telecom, Telecom SudParis, France
Dr. Dongkyun Kim, KISTI(Korea Institute of Science and Technology Information), Korea
Prof. Yoonhee Kim, Sookmyung Women s University, Korea
Prof. Li-Der Chou, National Central University, Taiwan
Prof. Young Woong Ko, Hallym University, Korea
Prof. Dimiter G. Velev, UNWE(University of National and World Economy), Bulgaria
Dr. Tadasuke Minagawa, Meiji University, Japan
Prof. Jun-Kyun Choi, KAIST (Univ.), Korea
Dr. Brownson ObaridoaObele, Hyundai Mobis Multimedia R&D Lab , Korea
Prof. Anisha Lal, VIT university, India
Dr. kyeong kang, University of technology sydney, faculty of engineering and IT , Australia
Prof. Chwen-Yea Lin, Tatung Institute of Commerce and Technology, Taiwan
Dr. Ting Peng, Chang'an University, China
Prof. ChaeSoo Kim, Donga University in Korea, Korea
Prof. kirankumar M. joshi, m.s.uni.of baroda, India
Dr. Chin-Feng Lin, National Taiwan Ocean University, Taiwan
Dr. Chang-shin Chung, TTA(Telecommunications Technology Association), Korea

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Dr. Che-Sheng Chiu, Chunghwa Telecom Laboratories, Taiwan
Dr. Chirawat Kotchasarn, RMUTT, Thailand
Dr. Fateme Khalili, K.N.Toosi. University of Technology, Iran
Dr. Izzeldin Ibrahim Mohamed Abdelaziz, Universiti Teknologi Malaysia , Malaysia
Dr. Kamrul Hasan Talukder, Khulna University, Bangladesh
Prof. HwaSung Kim, Kwangwoon University, Korea
Prof. Jongsub Moon, CIST, Korea University, Korea
Prof. Juinn-Horng Deng, Yuan Ze University, Taiwan
Dr. Yen-Wen Lin, National Taichung University, Taiwan
Prof. Junhui Zhao, Beijing Jiaotong University, China
Dr. JaeGwan Kim, SamsungThales co, Korea
Prof. Davar PISHVA, Ph.D., Asia Pacific University, Japan
Ms. Hela Mliki, National School of Engineers of Sfax, Tunisia
Prof. Amirmansour Nabavinejad, Ph.D., Sepahan Institute of Higher Education, Iran

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Editor Guide

■ Introduction for Editor or Reviewer

All the editor group members are to be assigned as a evaluator(editor or reviewer) to submitted journal papers at the discretion of the Editor-in-Chief. It will be informed by eMail with a Member Login ID and Password.

Once logined the Website via the Member Login menu in left as a evaluator, you can find out the paper assigned to you. You can evaluate it there. All the results of the evaluation are supposed to be shown in the Author Homepage in the real time manner. You can also enter the Author Homepage assigned to you by the Paper ID and the author's eMail address shown in your Evaluation Webpage. In the Author Homepage, you can communicate each other efficiently under the peer review policy. Please don't miss it!

All the editor group members are supposed to be candidates of a part of the editorial board, depending on their contribution which comes from history of ICACT TACT as an active evaluator. Because the main contribution comes from sincere paper reviewing role.

■ Role of the Editor

The editor's primary responsibilities are to conduct the peer review process, and check the final camera-ready manuscripts for any technical, grammatical or typographical errors.

As a member of the editorial board of the publication, the editor is responsible for ensuring that the publication maintains the highest quality while adhering to the publication policies and procedures of the ICACT TACT(Transactions on the Advanced Communications Technology).

For each paper that the editor-in-chief gets assigned, the Secretariat of ICACT Journal will send the editor an eMail requesting the review process of the paper.

The editor is responsible to make a decision on an "accept", "reject", or "revision" to the Editor-in-Chief via the Evaluation Webpage that can be shown in the Author Homepage also.

■ Deadlines for Regular Review

Editor-in-Chief will assign a evalaution group(a Editor and 2 reviewers) in a week upon receiving a completed Journal paper submission. Evaluators are given 2 weeks to review the paper. Editors are given a week to submit a recommendation to the Editor-in-Chief via the evaluation Webpage, once all or enough of the reviews have come in. In revision case, authors have a maximum of a month to submit their revised manuscripts. The deadlines for the regular review process are as follows:

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Evalution Procedure	Deadline
Selection of Evaluation Group	1 week
Review processing	2 weeks
Editor's recommendation	1 week
Final Decision Noticing	1 week

■ Making Decisions on Manuscript

Editor will make a decision on the disposition of the manuscript, based on remarks of the reviewers. The editor's recommendation must be well justified and explained in detail. In cases where the revision is requested, these should be clearly indicated and explained. The editor must then promptly convey this decision to the author. The author may contact the editor if instructions regarding amendments to the manuscript are unclear. All these actions could be done via the evaluation system in this Website. The guidelines of decisions for publication are as follows:

Decision	Description
Accept	An accept decision means that an editor is accepting the paper with no further modifications. The paper will not be seen again by the editor or by the reviewers.
Reject	The manuscript is not suitable for the ICACT TACT publication.
Revision	The paper is conditionally accepted with some requirements. A revision means that the paper should go back to the original reviewers for a second round of reviews. We strongly discourage editors from making a decision based on their own review of the manuscript if a revision had been previously required.

■ Role of the Reviewer

Reviewer Webpage:

Once logined the Member Login menu in left, you can find out papers assigned to you. You can also login the Author Homepage assigned to you with the paper ID and author's eMail address. In there you can communicate each other via a Communication Channel Box.

Quick Review Required:

You are given 2 weeks for the first round of review and 1 week for the second round of review. You must agree that time is so important for the rapidly changing IT technologies and applications trend. Please respect the deadline. Authors undoubtedly appreciate your quick review.

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Anonymity:

Do not identify yourself or your organization within the review text.

Review:

Reviewer will perform the paper review based on the main criteria provided below. Please provide detailed public comments for each criterion, also available to the author.

- How this manuscript advances this field of research and/or contributes something new to the literature?
- Relevance of this manuscript to the readers of TACT?
- Is the manuscript technically sound?
- Is the paper clearly written and well organized?
- Are all figures and tables appropriately provided and are their resolution good quality?
- Does the introduction state the objectives of the manuscript encouraging the reader to read on?
- Are the references relevant and complete?

Supply missing references:

Please supply any information that you think will be useful to the author in revision for enhancing quality of the paperor for convincing him/her of the mistakes.

Review Comments:

If you find any already known results related to the manuscript, please give references to earlier papers which contain these or similar results. If the reasoning is incorrect or ambiguous, please indicate specifically where and why. If you would like to suggest that the paper be rewritten, give specific suggestions regarding which parts of the paper should be deleted, added or modified, and please indicate how.

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Journal Procedure

Dear Author,

- **You can see all your paper information & progress.**

- **Step 1. Journal Full Paper Submission**

Using the Submit button, submit your journal paper through ICACT Website, then you will get new paper ID of your journal, and send your journal Paper ID to the Secretariat@icact.org for the review and editorial processing. Once you got your Journal paper ID, never submit again! Journal Paper/CRF Template

- **Step 2. Full Paper Review**

Using the evaluation system in the ICACT Website, the editor, reviewer and author can communicate each other for the good quality publication. It may take about 1 month.

- **Step 3. Acceptance Notification**

It officially informs acceptance, revision, or reject of submitted full paper after the full paper review process.

Status	Action
Acceptance	Go to next Step.
Revision	Re-submit Full Paper within 1 month after Revision Notification.
Reject	Drop everything.

- **Step 4. Payment Registration**

So far it's free of charge in case of the journal promotion paper from the registered ICACT conference paper! But you have to regist it, because you need your Journal Paper Registration ID for submission of the final CRF manuscripts in the next step's process. Once you get your Registration ID, send it to Secretariat@icact.org for further process.

- **Step 5. Camera Ready Form (CRF) Manuscripts Submission**

After you have received the confirmation notice from secretariat of ICACT, and then you are allowed to submit the final CRF manuscripts in PDF file form, the full paper and the Copyright Transfer Agreement. Journal Paper Template, Copyright Form Template, BioAbstract Template,

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Journal Submission Guide

All the Out-Standing ICACT conference papers have been invited to this "ICACT Transactions on the Advanced Communications Technology" Journal, and also welcome all the authors whose conference paper has been accepted by the ICACT Technical Program Committee, if you could extend new contents at least 30% more than pure content of your conference paper. Journal paper must be followed to ensure full compliance with the IEEE Journal Template Form attached on this page.

➢ How to submit your Journal paper and check the progress?

Step 1. Submit	Using the Submit button, submit your journal paper through ICACT Website, then you will get new paper ID of your journal, and send your journal Paper ID to the Secretariat@icact.org for the review and editorial processing. Once you got your Journal paper ID, never submit again! Using the Update button, you can change any information of journal paper related or upload new full journal paper.
Step 2. Confirm	Secretariat is supposed to confirm all the necessary conditions of your journal paper to make it ready to review. In case of promotion from the conference paper to Journal paper, send us all the .DOC(or Latex) files of your ICACT conference paper and journal paper to evaluate the difference of the pure contents in between at least 30% more to avoid the self replication violation under scrutiny. The pure content does not include any reference list, acknowledgement, Appendix and author biography information.
Step 3. Review	Upon completing the confirmation, it gets started the review process thru the Editor & Reviewer Guideline. Whenever you visit the Author Homepage, you can check the progress status of your paper there from start to end like this, " Confirm OK! -> Gets started the review process -> ...", in the Review Status column. Please don't miss it!

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Volume. 7 Issue. 5

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Detecting Anomalous Network Traffic in IoT Networks

Dang Hai Hoang*, Ha Duong Nguyen**

*Posts and Telecommunication Institute of Technology, Hanoi, Vietnam

**Faculty of Information Technology, National University of Civil Engineering, Hanoi, Vietnam

hdhai.hn@gmail.com, nghaduong@gmai.com

***Abstract*—Network operators need effective tools to quickly detect anomalies in traffic data for identifying network attacks. In contrast to traditional Internet, detection of anomalous network traffic in IoT (Internet of Things) networks is becoming a challenge task due to limited network resources and performance. Comprehensive detection methods are no longer effective for IoT networks, calling for developing lightweight solutions. Principal Component Analysis (PCA) techniques can help to reduce computing complexity, thus, anomaly detection techniques based on PCA received a lot of attention in the past. However, PCA techniques could not be directly applied to IoT networks with constrained resources and limited performance. This paper investigates PCA techniques for detecting anomalous network traffic in IoT networks. We propose a novel detection scheme with two levels using PCA techniques. The first level is for quick detection with few principal components while the second level is for detailed detection with a number of principal components. We investigate the selection of parameters in a distance calculation formula using several experiments to show the feasibility of our proposed scheme.**

***Keyword*— IoT Network Traffic Anomaly, Anomaly Detection, Principal Component Analysis, Information Security, Network Security**

I. INTRODUCTION

THE world of interconnected things - the Internet of Things (IoT) opens a number of challenges regarding security. IoT can bring huge interested services nowadays, but there is still a lack of suitable security measures for mixed IoT environment [1]. IoT devices have constrained resources and limited performance and are attractive to the attackers [2,3]. Cyber attacks in IoT networks are becoming more difficult to detect. Traditional mechanisms are usually comprehensive and no longer effective for IoT networks, calling for developing new paradigm. Given the growing complexity of connected things and many constraints of IoT environment, there is a strong demand for developing new effective security solutions.

Network traffic anomaly detection (NTAD) is a promised approach for identifying network attacks since it can detect new attacks without pre-recorded signature. That's why NTAD received a lot of attention in recent years [4,5]. The definition of anomaly was formally given by Hawkins as "an observation which deviates so much from other observations as to arouse uncertainties that it was produced by an alternative mechanism" [4]. Network anomalies are unusual patterns in traffic data that do not conform to expected normal behaviour. Anomalies may be performance related (due to network failures, changes in link traffic, flash crowd, etc...) or security related (due to attacks such as denial of service attacks, network scans, etc…). NTAD is a very critical task of network operators. Effective tools are necessary for quick detection of exceptional non-conforming patterns in traffic data in order to identify abnormal traffic flows or the causes of anomalies for further handling.

Many anomaly detection techniques have been proposed in the past within diverse research areas and application domains, see e.g. [4-11]. Traditional techniques considered anomalies as outlier and typically proposed to use statistical properties of observed data to construct a normal profile based on normal traffic data. The current collected data will be compared to this profile for checking any deviation to detect anomalies [5-8, 11]. General speaking, the main principle of NTAD is to build the baseline (the normal region) using features of network traffic in normal condition and to compare online collected traffic data to this baseline to find out deviation (anomalies). However, this process is difficult in practice due to many factors such as: multivariate features of traffic data, correlation of various data features, complex dataset, required accuracy and detection speed, etc.

On the other hand, NTAD for IoT networks is facing other issues such as: heterogeneity, constrained resources and limited performance of IoT devices. In IoT networks, we can not use complex methods as in traditional Internet. Lightweight techniques with less complexity, low resource usage and lower computation requirement are necessary. The development of such techniques remains a challenging research task.

Among broad existing anomaly detection techniques [4-11] for the Internet, multivariate statistical approaches for anomaly detection received a lot of attention. Principal Component Analysis (PCA) is one of the best-known multivariate statistical analysis techniques for detecting anomalies in the context of constrained network resources like IoT networks. PCA is a dimension reduction technique, which transforms a set of correlated original variables into a set of few uncorrelated variables, called Principal Components (PC). These PCs are linear combinations of the original variables. The number of PCs is less than or equal to

Manuscript received on December 17th, 2017. This work is follow-up of the invited journal to an accepted paper of the 20th International Conference on Advance Communication Technology (ICACT2018). The work is supported by the ASEAN IVO Project "A Hybrid Security Framework for IoT Networks".

Dang Hai Hoang is with the Posts and Telecommunication Institute of Technology, Hanoi, Vietnam (Corresponding author to provide phone: +84-4- 3854-4451; fax: +84-4-3756-2036; e-mail: haihd@ptit.edu.vn).

Ha Duong Nguyen is with the National University of Civil Engineering, Hanoi, Vietnam (e-mail: nghaduong@gmail.com).

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

the number of original variables. Thus, PCA allows lower complexity. PCA is considered as a simple but effective method for NTAD [5, 9-19].

Using PCA for NTAD is an attractive approach presented in pioneer research works by Shyu et al. [12], Lakhina et al. [13], Brauchkoff et al. [14], Ringberg et al. [15] and Kwitt et al. [19]. A variety number of further works has been proposed based on PCA with several enhancements such as [9-11, 16-29]. However, several issues in applying PCA have still not been considered as described in [4, 5, 7, 10, 16] including sensitivity, effectiveness, dimension-depending computation complexity, feature selection. Two main issues for the goodness of PCA based anomaly detection are: how many PCs are to select and how to calculate the distance (the deviation measure).

The issue of selecting PCs has been investigated in several works including the selection of major PCs and minor PCs [11, 12, 18, 20, 23], the selection of PC subspace [14-18, 28-30]. However, heuristic is the common way of choosing PCs. It is not clear, how many PCs and which PCs are to select. On the other hand, the choice of distance formula is not clear in most of the research works. Distance is a quantitative metric for deviation of traffic pattern. Most of the proposed methods are using either T^2 distribution formula [4-6, 10-13], or Euclidean distance [7, 11, 23, 26], or Mahalanobis distance [7, 10-18, 20-22, 24-29]. It is not clear how the choice of the distance formula will affect the detection accuracy and the computation complexity. On the other hand, the high complexity of such formulas is not suitable for a quick online detection of traffic anomalies. Such problems will have impact on further development network traffic anomaly detection based on PCA.

This paper addresses the problems for detection of network traffic anomaly in the context of IoT networks with resource constraints. We investigate PCA techniques using a new general formula for deviation (distance) calculation. We show that distance formulas used in previous typical research works can be derived from our general formula. To our best knowledge, the proposed formula is the first one for interpreting the parameters in different distance formulas. Based on selecting PCs using this general formula, we propose a novel detection scheme with two levels using PCA techniques. The first level is for quick detection with few k principal components in order to reduce the complexity to $O(k)$ while retaining acceptable detection results in comparison to previous methods. The second level is for detailed detection with a number of principal components.

The rest of this paper is organized as follows. Section II presents the basic PCA techniques and related works. concept of PCA and anomaly detection. Section III describes related works. Section IV proposes a new distance formula and our PCA scheme. Section V presents the experiments. Section VI concludes the paper.

II. PCA Techniques and Related Works

A. The Basic Concept of PCA

Anomaly detection often requires a high dimension data including many features (attributes) collected from networks. Therefore, the analysis has high computation complexity, needs much time, and is not suitable for quick detection requirement. PCA is the most common technique to reduce high dimension of data [4-8]. It converts the original data into new set of axes called principal components (PC) by keeping the most essential features of the original data.

Let X be the original observed dataset with n rows and p columns $\{X_1, X_2, \ldots, X_p\}$. The dataset is a $n \times p$ matrix, each column represents an attribute (feature) of data. Each column is represented by a p-dimensional vector of p correlated variables. Let R be a pxp sample correlation matrix of $\{X_1, X_2, \ldots, X_p\}$. Let $(\lambda_1, \boldsymbol{e}_1)$, $(\lambda_2, \boldsymbol{e}_2)$, $(\lambda_3, \boldsymbol{e}_3)$,… $(\lambda_p, \boldsymbol{e}_p)$ be p eigenvalue and eigenvector pairs of the matrix R. PCA converts X into a new dataset Y using transformation matrix R as follows:

$$\boldsymbol{Y = RX} \quad (1)$$

We have the i^{th} ***PC*** as follows:

$$y_i = \boldsymbol{e}_i^T (\boldsymbol{x} - \bar{\boldsymbol{x}}), \quad (2)$$

where $i = 1, 2, \ldots, p$

$\boldsymbol{e}_i = (e_{i1}, e_{i2}, \ldots e_{ip})^T$ is the i^{th} eigenvector.

x is the observation

$\bar{x} = (\sum_{i=1}^{n} x_i)/n$ is the sample mean of x.

The eigenvalues (λ) are the roots of equation

$| R - \lambda I | = 0$.

Each eigenvalue has a corresponding non-zero eigenvector $\boldsymbol{e}$, which satisfies: $R\boldsymbol{e} = \lambda \boldsymbol{e}$

Let $\boldsymbol{z}=(z_1, z_2, \ldots z_p)'$ be the vector of standardized observations, i.e. $\boldsymbol{z} = \boldsymbol{x} - \bar{\boldsymbol{x}}$, we rewrite (2) as:

$$y_i = \boldsymbol{e}_i^T \boldsymbol{z} \quad (3)$$

PCA has the following important properties. The PCs are uncorrelated. The first PC has the highest variance; the second PC has the next highest variance, and so on. The total variance in all PCs combined is equal to the total variance of the original variables $X_1, X_2, \ldots, X_p$. The PCs are sorted in descending order of the eigenvalues, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots \geq \lambda_p > 0$. The quotient $\lambda_i / \sum_{j=1}^{p} \lambda_j$ describes the contribution of the i^{th} PC on the variance of the data. The dataset X is transformed to PCA based on eigenvectors with the target that the produced PCs have the highest possible variances.

B. Anomaly Detection Using PCA

The common principle for anomaly detection using PCA is to calculate the statistical distance from each observed data to the normal dataset (i.e. to the centroid or statistical average of the dataset). The distance calculation is performed in the principal component space. Mathematically, distance is a quantitative degree of how far two data instances apart from each other. Observed data instances which are at far distance from the new axes represented by PCs are considered as abnormal behaviour. For the comparison, a threshold value is established. If the calculated distance of the observed data is larger than the threshold, this data instance is considered as an anomaly. The threshold value is usually determined by the statistical distribution function of the distance [4-6, 30].

The most popular distance formulas used in previous anomaly detection methods are the Euclidean distance, the Mahalanobis distance or the statistic T^2 distribution formula. We use the following notations in the formulas.

Let $\mathbf{x} = (x_1, x_2, \ldots, x_p)$ and $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)$ be the observed data, where p is the number of original attributes, and is the number of input variables.

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Let $d(\boldsymbol{x}, \boldsymbol{\mu})$ is the distance between two points $\boldsymbol{x}$ and $\boldsymbol{\mu}$ in the PCA space.

Let d_N is the threshold, which is calculated using the normal data profile as presented in [5-8,11].

1) *The Euclidean Distance*

Intuitively, the most common distance function is the Euclidean distance. The Euclidean distance between **x** and **μ** is [7, 11]:

$$d(\boldsymbol{x}, \mu) = \sqrt{(\boldsymbol{x}-\mu)^T(\boldsymbol{x}-\mu)} \quad (4)$$

If $\boldsymbol{\mu}$ is the vector of mean values of each input variable, the Euclidean distance in PCA space is given by:

$$\Delta_{Euclid} = \sqrt{\sum_{i=1}^{p} y_i^2} \quad (5)$$

The data instance $\boldsymbol{x}$ can be determined as anomalous if $\Delta_{Euclid} \geq d_N$

In order to reduce the square root computation complexity, the square distance is also usually used:

$$d_E = \Delta_{Euclid}^2 = \sum_{i=1}^{p} y_i^2 \quad (6)$$

The Euclidean distance formula is simple, but does not take into account the variance of each variable.

2) *The Mahalanobis Distance*

The Mahalanobis distance between x and μ is computed as [7, 10-12, 14-18, 20-22, 25-29]:

$$d(\boldsymbol{x}, \mu) = \sqrt{(\boldsymbol{x}-\mu)^T \boldsymbol{S}^{-1}(\boldsymbol{x}-\mu)} \quad (7)$$

Where S is the covariance matrix, S^{-1} is the transposition of the sample covariance matrix between $\boldsymbol{x}$ and $\boldsymbol{\mu}$. The covariance matrix is used as weights to reduce the different variance between the variables. Thus, the covariance matrix represents the relationship between variables more efficient.

In PCA space, the Mahalanobis distance is computed as:

$$\Delta_{Ma} = \sqrt{\sum_{i=1}^{p} \frac{y_i^2}{\lambda_i}} \quad (8)$$

The data instance $\boldsymbol{x}$ can be determined as anomalous if $\Delta_{Ma} \geq d_N$

The square distance is usually used as

$$d_m = \Delta_{Ma}^2 = \sum_{i=1}^{p} \frac{y_i^2}{\lambda_i} \quad (9)$$

The Mahalanobis distance formula takes into account not only the average value but also the variance and covariance of the variables. Instead of simply computing the distance from the mean value, it weights each variable by its standard deviation and covariance [25].

3) *Hotelling's T2 Statistic*

Another way to calculate the distance is using statistical distribution, typically the Hotelling's T^2 statistic [10-13,30]. This statistic gives the deviation (distance) from each vector **x** to the centroid **μ** of the statistical distribution of the data.

$$T^2 = (\boldsymbol{x}-\bar{\boldsymbol{x}})^T \boldsymbol{S}^{-1}(\boldsymbol{x}-\bar{\boldsymbol{x}}) \quad (10)$$

Where S^{-1} is the transposition of the sample covariance matrix between $\boldsymbol{x}$ and the sample mean of $\boldsymbol{x}$.

In PCA space, the statistic T^2 distribution is computed as:

$$T^2 = \sum_{i=1}^{p} \frac{y_i^2}{\lambda_i} \quad (11)$$

In comparison to (9), the statistic T^2 is similar to the Mahalanobis distance.

T^2 is distributed as $\frac{(n-1)p}{n-p} F_{p,n-p}$, where $F_{p,n-p}$ denotes a random variable with an *F*-distribution with *p* and *n-p* degrees of freedom. A large value of T^2 indicates a deviation of the observation $\boldsymbol{x}$ from the centroid of the normal dataset and the F-statistic can be used to test for an anomaly [12, 30].

C. *Related Works on NTAD using PCA*

The research works on NTAD using PCA were initiated by Shyu et al. [12] and Lakhina et al. [13]. In [12], the authors proposed a principal component classifier (PCC) consisting of two functions of PC scores, one for major components (Major PCs) and one for minor components (Minor PCs). The major PCs are used to detect extreme observations with large values on some original variables. The minor PCs are used to detect observations that do not conform to the normal correlation structure. The Mahalanobis distance is used with two thresholds, one for major PCs and another for minor PCs. The detection rate is depending on the quality and the accuracy of PCC. Thus, some other authors [20, 21] proposed to improve PCC with multivariate trimming for robustness, threshold calculation.

Lakhina et al. [12] proposed to divide the network traffic data into normal subspace consisting of typical behaviour pattern and anomalous subspace accounting for uncharacteristic circumstances by mean of PCA. The Euclidean distance is employed to check the dissimilarity between data instances. Anomaly detection is using the first PCs and the mean value of the distance. However, there are several remaining issues to be solved such as the choice of PCs, the separation of normal and anomalous subspaces. Further improving approaches were indicated in [17, 26] with the concept of traffic matrix including a number of original/ destination flows (OD flows). The authors proposed to extract a list of few PCs, which contain the maximum variance of the original data. The authors in [16] proposed a control approach for pre-processing the network data to be analysed with PCA. The authors in [18] proposed to decompose the traffic variations into normal and anomalous components. A new method for identifying the anomalous flows inside the aggregated flows was introduced.

The works of Shyu et al [11] and Lakhina et al. [12] received a lot of attention in the research community. The authors in [14] presented the issue of temporal correlation in previous works of Shyu and Lakhina, and proposed a predictive filter for classical PCA in order to improve correlation effect. Ringberg et al. [15] indicated that PCA techniques are very sensitive for traffic anomaly detection, especially for the selected parameters. The authors showed that current methods for tuning PCA parameters are inadequate and presented several challenges of using PCA. However, they concluded that PCA is a promising statistical analysis technique that can be used effectively for detecting network anomalies.

In recent years, a variety number of research works has been developed based on the application of PCA for anomalous traffic detection. The authors in [9] presented the application of PCA subspace method for anomaly detection in backbone networks. PCA is used for investigate the explored data with a lower approximation. Zargar et al [22] discussed category-based selection of the features for PCA based anomaly detection. By classifying network traffic into six typical groups, the paper proposed to select most important features in order to reduce the amount of data to be

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

analyzed by PCA. Huang et al. [26] presented three major approaches for network traffic anomaly detection: PCA-based, sketch-based and signal-analysis-based. The authors tried to combine these approaches into a unified frame. Lee at al. [23] proposed an online over-sampling PCA algorithm for anomaly detection. Unlike other PCA based approaches, this method does not store the entire data matrix. Instead, it uses an online updating technique to update the principal direction without solving eigenvalue decomposition problems. The authors claimed that this method is favored for online applications. Camacho et al. [17] proposed a sketch-based algorithm in addition to traditional PCA for traffic anomaly detection. In this algorithm, each local monitor only maintains a series of sketches for each traffic flow in order to reduce the raw data. The anomaly detection is following the method proposed by Lakhina [12] with the novel concept of residual subspace. Horrou et al. [10] proposed an integration of PCA, Hotelling's T^2 and Q statistics to detect small or moderate anomalies in the process mean. The authors showed that T^2 statistic can result in false negatives (missed detection) and cannot detect anomalies that are orthogonal to the first PCs.

Anomalies can be considered as outliers. The authors in [6] provided a survey on outlier detection techniques and applications. The authors in [11] discussed three statistical techniques in intrusion detection: PCA-based, Chi-square distribution and Gaussian mixture distribution and discussed the comparative performance of them. The authors in [19] presented a multi-step outlier-based approach for anomaly detection in network-wide traffic. A subset of dataset is called a cluster consisting of similar data objects.

The choice of a reasonable number of PCs is important for alleviate the autocorrelation of the residual PCs. The authors [7, 10, 25-29] concluded that the goodness of the PCA model depends on a good choice of how many PCs are retained. The paper [10, 28] indicated that techniques such as Scree plot and cumulative percentage variance (CPV) can be used to determine the number of PCs for the PCA model. Nevertheless, a threshold value of cumulative variance is needed to determine the number of PCs to use. Bhuyan et al. [8] presented an overview of various facets of network anomaly detection. The paper provided a broad survey of the existing research on network anomaly detection methods in the context of network security.

Anomalies are usually detected by the help of a distance measure [19]. A good survey of distance measures used within network anomaly detection was given in [7]. The paper discussed the theoretical background in distance measures and various types of distance measures published in previous papers. The authors discussed two common distance measures including Euclidean distance [7, 11, 22, 24-26] and Mahalanobis distance [10-12, 14-21, 25-29], and several equivalent distance measures such as weighted Euclidean distance, Manhattan distance, distribution law distance T_2. The paper [7] also indicated the limitation on identifying and selecting distance measures for network anomaly detection. The paper [29] discussed the use of Mahalanobis distance for anomaly detection in time series in the context of an unsupervised learning algorithm. Fan et al. [27] proposed a modified PCA scheme which uses multiple similarity measurements to generate multiple subspaces. The similarity is measured by the Mahalanobis distance. Unlike other PCA based subspace methods that use only one measurement, MPCA uses multiple measurements. This scheme has the main disadvantage of computation complexity. It has proposed mainly for learning approaches, in particular for image processing, and it is not suitable for quick detection of anomalous network traffic. In [8], the appropriateness of proximity measures (distance measures) was discussed, but not in details. A summary of distance measures was given but there is no explanation for the choice of measures. Bayarjargal et al. [25] proposed a combination of entropy distribution and Mahalanobis distance for anomaly detection. First, entropy of selected attributes is computed for defining suspicious traffic area. After that, Mahalanobis distance is calculated to check anomalies.

D. Issues of Previous PCA Based Methods

PCA allows dimensionality reduction, thus, it is suitable to handle high dimensional data sets. It converts a multivariate high dimensional data into uncorrelated individual PCs. In principle, tests for anomaly can be applied on individual PCs. However, standard PCA based techniques are typically proportional quadratic in the number of variables regarding the distance calculation. The complexity is typically $O(kn^2)$ with k is the number of PCs and n is the number of original variables [4, 17, 23, 26]. Thus, many research works tried using a subset of PCs in order to reduce the computation complexity [6-11, 23, 26-29].

As presented in section II.A, the PCA transformation is performed in such a way that the first PCs account for the most of the variance in the original data, and the last PCs represent linear functions of the original variables with very small variance. The first PCs represent the large cumulative proportion of the total variance of the original data. Thus, these PCs tend to be strongly related to the anomalies on one or more original variables. Intuitively, it is reasonable to detect anomalies by looking at few first PCs [18]. If most of the variance of an n-dimensional data set is accounted by $k < n$ principal components, it is possible to select only first k PCs. The dimension of the data is then reduced to k [26]. However, it is still not clear how many PCs to be selected for an effective detection.

On the other hand, last PCs are sensitive to the observation that are inconsistent with the correlation structure of the data as indicated in several works such as [12-16]. Large values on few last PCs will reflect multivariate anomalies that are not detectable using the criterion based on large values of the original variables [10]. It is useful to check the last PCs for a significant variation of an observation. That is why many works suggested using few first PCs and few last PCs.

Typically, two thresholds are used: one for major PCs (q first PCs) and one for minor PCs (last r PCs). As indicated in [12, 17, 20, 21], a data instance $\boldsymbol{x}$ is anomalous if

$$\sum_{i=1}^{q} \frac{y_i^2}{\lambda_i} > c_1 \text{ or } \sum_{i=p-r+1}^{p} \frac{y_i^2}{\lambda_i} > c_2 \quad (12)$$

and is normal if

$$\sum_{i=1}^{q} \frac{y_i^2}{\lambda_i} \le c_1 \text{ and } \sum_{i=p-r+1}^{p} \frac{y_i^2}{\lambda_i} \le c_2 \quad (13)$$

Where c_1 and c_2 are the thresholds, which are calculated with the help of the chi-square distribution. These thresholds

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

are very sensitive to the detection results.

Several works followed the subspace approach and proposed the decomposition of the data into normal and anomalous subspaces using projection of data on the first PCs and the last PCs respectively [13, 24, 26]. The subspace methods are also called residual analysis methods. One issue with this approach is that PCA based anomaly detection techniques are sensitive to the number of PCs in each of the two subspaces [5, 7, 16].

The choice of distance formula is one issue of most of the PCA based anomaly detection methods. As presented in the above section, most of the proposed methods are using either T^2 formula [4-6, 30], or Euclidean distance [7, 17, 24, 26] or Mahalanobis distance [10, 11, 16, 20, 21, 25]. The computation of the distance is quadratic proportional to of the variables. With high dimensional dataset, the distance calculation will increase the computational cost, and a quick (online) detection is not possible.

The application of PCA techniques for NTAD in IoT networks has been just investigated in few recent research works, e.g. [31-37]. A general security framework was recently proposed in [31], which is based on three parts of IoT systems, namely physical part, network part and application part. However, there is still not clear how to apply detection methods for network traffic anomalies. The paper just described the general scheme, which is mainly for authentication. The paper [32] provided basic three layered IoT architecture and discussed several security issues of IoT that exist in such architectures. There is a good survey on previous research works on IoT security. The authors in [33] provided a survey of several security architectures for embedded IoT systems. These architectures are mainly based on hardware-specific environment and are typically designed for specific tasks. Pajouh et al. [34] investigated the demand on detecting intrusion and malicious activities within IoT networks and proposed a model for intrusion detection based on two-layer dimension reduction and two-tier classification module. The model is using PCA and linear discriminate analysis. Sharma et al. [35] proposed a framework for coordinated processing between edge and cloud computing / processing. The paper discussed the issue of hue amount of data generated from heterogeneous wireless IoT devices. Ferrando et al. [36] investigated the issue of streaming analytical techniques for detecting events in traffic feature distribution, which can allow the classification of abnormal behaviour within an IoT network. Recently, Zhao et al. [37] proposed a model for intrusion detection based on dimension reduction algorithm using PCA and a classifier for IoT networks.

In fact, in contrast to the traditional networks, IoT opens a completely new dimension to security, where attack threats move from manipulating information to controlling actuation (in other words, moving from the digital to the physical world). In our best knowledge, there is still very few works investigating PCA based methods for NTAD in IoT networks. Moreover, there are still several issues for network traffic anomaly detection using PCA, which have been not adequately investigated for IoT environment regarding the constrained resource and limited performance. Three challenges are of most interested by applying PCA for IoT networks including: 1) selection of PCs for an effective detection regarding the IoT network constraints, 2) distance measure for lower complexity and 3) quick detection of network traffic anomalies. These issues are the topics of this paper.

III. A Novel Network Anomaly Detection Scheme using PCA for IoT Networks

A. Overview of the Proposed Concept

Figure 1 describes the network model we use for our network anomaly detection scheme. The concept of fog was introduced. A fog is a network architecture for processing data and events from IoT devices closer to the sources of data in contrast to the central data network (known as "Cloud infrastructure". In other words, fog architecture extends the "cloud" (the cyber world) into the physical world of things (the world of IoT devices). Within the network model, the network traffic anomaly detection (NTAD) can be implemented into two levels. NTAD Level 1 (NTAD-L1) is for the fog architecture (IoT network segment closed to devices) with constrained resources and performance. This detection level requires lower complexity. NTAD Level 2 is for the cloud infrastructure with powerful resources and computing performance. NTAD Level 1 uses PCA technique with few PCs for quick detection, while NTAD Level 2 can deploy more PCs for detailed detection. One NTAD Level 2 module can serve several NTAD Level 1 modules (e.g. NTAD-L1-1, NTAD-L1-2, NTAD-L1-3).

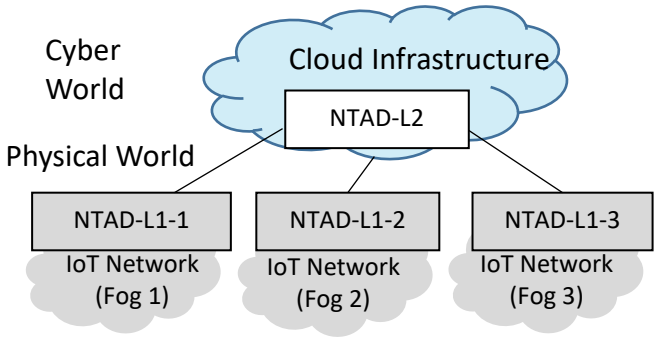


Fig. 1. The Network Model.

Figure 2 depicts the general flow of our anomaly detection scheme using two levels. Network traffic is collected from the IoT network (the fog) by a flow capture (like in other works such as [24]).

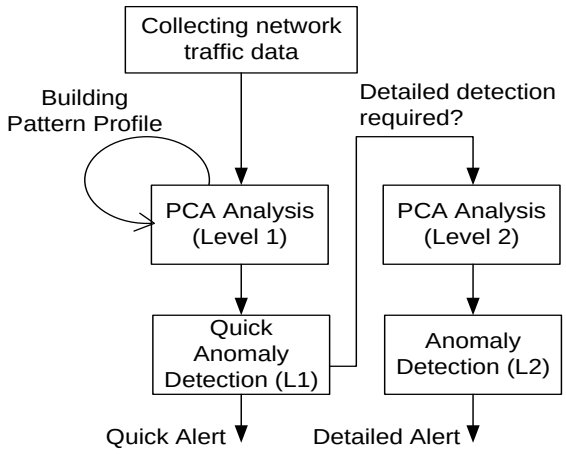


Fig. 2. The Detection Flow.

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

The captured data is standardized and transformed into PCA. Based on traffic data in normal condition (no attacks), the model builds a pattern profile (the baseline) for comparison later. The Level 1 detection module uses few PCs (e.g. 3 PCs) for quick anomaly detection. The number of suitable PCs will be given in the next section. If the network operator wants to have more detailed detection alerts, he can proceed the Level 2 detection module, which uses more PCs (e.g. major PCs and minor PCs as presented in [12, 17-22, 24-26]) for providing detailed detection.

In the following section, we present our detection scheme in details.

B. A Novel Distance Formula for the Detection Scheme

As presented in Section II, popular NTAD methods use Euclidean distance or Mahalanobis distance or statistic T^2 distribution with high computing complexity. Thus, such methods are not suitable for online anomaly detection, especially for IoT networks (the fogs).

For quick online anomaly detection, we suggest to use only few PCs and a simple distance formula as possible. Intuitively, the price to be paid is the accuracy. However, the detection can be implemented into two steps as described above. At the first step, network operators just want to know whether there are abnormal traffic flows or not. A quick detection with acceptable accuracy is desirable. At the next step, network operators can apply a detailed analysis for higher accuracy.

For implementing our proposed detection scheme with two levels, we want to develop a general distance calculation formula, which can use few PCs as well as a number of PCs. Our idea is to use the well-known Minkowski formula [38] for developing the new formula. The derivation of the general formula is as follows.

As indicated in [38], the Minkowski distance between two observed data $\boldsymbol{x} = (x_1, x_2, \ldots, x_p)$ and $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)$ is as follows:

$$\Delta_{Minkowski} = \left(\sum_{i=1}^{p} |x_i - \mu_i|^c \right)^{1/c}, c \geq 1; \qquad (14)$$

$$d = \Delta^c_{Minkowski} = \sum_{i=1}^{p} |x_i - \mu_i|^c, c \geq 1; \qquad (15)$$

By transforming into PCA domain with y_i principal components, we derive the Minkowski distance from each observation to the centre (the origin of PC axes) as follows:

$$d = \sum_{i=1}^{p} |y_i|^c, c \geq 1; \qquad (16)$$

Since each variable can have different variance, we can follow the approach of weighted Euclidean distance as illustrated in [16, 30] to give a weight to each PC in the equation (16) as follows:

$$d = \sum_{i=1}^{p} w_i |y_i|^c, \quad c \geq 1; \qquad (17)$$

Where w_i is the weight for the i^{th} PC ($w_i \neq 0$).

The formula (17) is our new proposed distance formula, which has the following properties.

Lemma 1: *The formula (17) is a general form of the Euclidean distance.*

Proof: Since c is arbitrary variable, we can choose c=2. All PCs can be considered equally, i.e. no variance of each input variable, the weight w_i can be set to 1. Thus, from (17) we can have:

$$d = \sum_{i=1}^{p} y_i^2 = y_i^2 + y_1^2 + \cdots + y_p^2 \qquad (18)$$

The formula (18) is the square Euclidean distance. In case we have consider variance of the input variables, i.e. PCs have different impacts on the distance calculation, we get:

$$d = \sum_{i=1}^{p} w_i y_i^2 \qquad (19)$$

The formula (19) is the weighted Euclidean distance. □

Lemma 2: *The formula (17) is a general form of the Mahalanobis distance.*

Proof: Since c is arbitrary variable, we can choose c=2. The impact of PCs is inverse proportional to their eigenvalues [38]. PCs with lower eigenvalues have large contribution on distance deviation, and PCs with large eigenvalues have less contribution on distance deviation, respectively. Thus, from (17) we can have:

$$d = \sum_{i=1}^{p} \frac{y_i^2}{\lambda_i} = \frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2} + \cdots + \frac{y_p^2}{\lambda_p} \qquad (20)$$

The formula (20) is the Mahalanobis distance. □

Lemma 3: *The formula (17) is a general form of the Mahattan distance.*

Proof: Since c is arbitrary variable, we can choose c=1. Thus, from (17) we can have:

$$d = \sum_{i=1}^{p} w_i |y_i| \qquad (21)$$

All PCs can be considered equally, i.e. no variance of each input variable, the weight w_i can be set to 1. Thus, from (17) we have:

$$d = \sum_{i=1}^{p} |y_i| = |y_1| + |y_2| + \cdots + |y_p| \qquad (22)$$

The formula (22) is the Mahalanobis distance. □

Conclusion:

We can use the general formula (17) for distance calculation in our proposed NTAD scheme based on PCA. For quick detection (Level 1 detection), we can use formula (21) or formula (22) with few PCs. In our previous works [28, 39], we have indicated that 3 or 4 PCs are enough for quick detection with an acceptable detection true rate of 92%. In this case, the computation complexity is O(*k*) with k=3 or 4.

For detailed detection (Level 2 detection), we can use more PCs using the formula (21) or (22), or even the formula (19) or (20), depending on the performance of the detection module NTAD-L2 in the cloud infrastructure.

C. Description of the Novel NTAD Scheme

In this section, we describe how to use the distance formula for our novel NTAD scheme.

At the initial stage, we have to collect the normal network

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

traffic (in case of no attacks) in order to build the pattern profile. The captured data is standardized and transformed into PCA dataset. By transforming into PCA space, we classify PCs into *m* major PCs and *p-m* minor PCs similar to the approach in [12, 17-22, 24- 27]. As suggested in [12, 26], the major PCs include the PCs representing 50% of the variance of total eigenvalues. The minor PCs have eigenvalues smaller than or equal to 0.2.

From the pattern profile, we can use the empirical cumulative distribution function (ECDF) to determine the thresholds d_{1N} for major PCs and d_{2N} for minor PCs, respectively. The ECDF function is defined as follows:

$$F_{ECDF} = P(d \le d_N) \quad (23)$$

That means the probability for a distance *d* smaller or equal to the threshold d_N. With the assumption of statistical distribution and an estimation rate α for false estimation, we can look at the table for F_{ECDF} - distribution function and seek for the values of d_N corresponding to (1-α) of the ECDF. For instance, if $\alpha = 5\%$, the thresholds d_N can be determined corresponding to *95%* of ECDF.

At the detection stage, online traffic data will be gathered into datasets. Similar to the previous stage, the collected dataset will be standardized and transformed into PCA domain using two subspaces, one for *m* major PCs and *p-m* minor PCs (see e.g. [12].

For quick detection (NTAD-L1), we apply the formula (22) with 3 or 4 PCs as follows:

$$d_{L1} = \sum_{i=1}^{k} |y_i| = |y_1| + |y_2| + \cdots + |y_k| \quad (24)$$

Where *k=3* or *k=4*.

The quick alert for traffic anomaly is determined if:

$$d_{L1} = \sum_{i=1}^{k} |y_i| > d_{NL1} \quad (25)$$

The threshold value d_{NL1} is determined using ECDF function as described above.

For detailed detection (NTAD-L2), we apply the formula (17) for *q* first PCs (*q*<*p*) including *m* major PCs and *q-r* minor PCs with *1*<*m*<*q-r*<*p*. An observed dataset is considered as anomaly if:

$$d_1 = \sum_{i=1}^{m} w_i |y_i|^c > d_{N1} \quad \text{OR} \quad (26)$$

$$d_2 = \sum_{i=r}^{q} w_i |y_i|^c > d_{N2} \quad (27)$$

Where d_1 and d_2 are the distances according to major PCs and minor PCs respectively, d_{1N} and d_{2N} are the corresponding thresholds of d_1 and d_2, OR and AND are the logic operations.

An observed dataset is considered as normal if:

$$d_1 = \sum_{i=1}^{m} w_i |y_i|^c \le d_{N1} \quad \text{AND} \quad (28)$$

$$d_2 = \sum_{i=r}^{q} w_i |y_i|^c \le d_{N2} \quad (29)$$

In case of *c=1* and w_i =1, the formulas for d_1 and d_2 will be:

$$d_1 = \sum_{i=1}^{m} |y_i| = |y_1| + |y_2| + \cdots + |y_m| \quad (30)$$

$$d_2 = \sum_{i=r}^{q} |y_i| = |y_r| + |y_{r+1}| + \cdots + |y_q| \quad (31)$$

The determination of the thresholds d_{1N} and d_{2N} is using the empirical cumulative distribution function (ECDF) as described above.

IV. Experiments

A. Dataset and Metrics

The dataset Kyoto Honeypot [40] is a real dataset collected from networks using a Honeypot at the Kyoto University. This dataset was used in many works for anomaly detection. The dataset includes most of the anomalies originated by the Internet. Thus, this dataset reflects objectively the anomalous events of collected network traffic.

We use in our experiments 14 essential features of the Kyoto Honeypot dataset including the most important features (attributes) of the network layer and transport layer traffic data. These features (attributes) are selected similar to the previous works [9, 12-21, 23-27] for comparison purpose.

For the experiments, we select three random datasets of the Kyoto Honeypot datasets collected from networks as indicated in the Table I.

TABLE I
THREE RANDOM DATASETS FOR EXPERIMENTS

Dataset	Number of flows	Number of anomalous flows	Number of normal flows
Dataset 1	125643	84476	41167
Dataset 2	114148	38790	75358
Dataset 3	120857	57337	63520

Following parameters are used for checking the accuracy:

$$TPR = TP / (TP + FN) \quad (32)$$

$$FPR = FP / (TN + FP) \quad (33)$$

Where *TPR* is True Positive Rate, *FPR* is False Positive Rate, *TP* is the number of true positive alerts, *FN* is the number of false negative alerts, *FP* is the number of false positive alerts, and *TN* is the number of true negative alerts. The total number of anomalous events is *TP+FN*; the total number of normal events is *TN+FP*.

B. Experiment Results

To build the sample dataset (the profile), we use traffic data from 5000 connections (similar to previous works in [12,13] for the consequent comparison). We transform this dataset into PCA space; calculate the centroid of the data using the distance formula. We determine the threshold using the empirical cumulative distribution function as described in the previous section. The thresholds are 95% of the ECDF function.

In the detection phase, the scheme uses the online collected datasets from Kyoto Honeypot datasets [40] and carries out the following steps: data standardization, PCA transformation, distance calculation, threshold comparison.

Table II shows the experiment result with dataset 1 using the threshold 95% for different combination of the

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

parameters k (number of PCs), *c* and w_i (three cases). As indicated in the grey row, our scheme provides acceptable TPR (94.3%) and FRP (4.8%) while keeping a low complexity with *k*=3 (only 3 PCs), c=1 and w_i =1.

TABLE II
RESULTS WITH DATASET 1, THRESHOLD 95%

Case	*k*	*c*	w_i	TPR (%)	FPR (%)
1	3	2	1	92.4	4.7
1	5	2	1	91.8	5.2
1	14	2	1	94.4	5.3
2	3	2	$1/\lambda_i$	94.6	4.9
2	5	2	$1/\lambda_i$	91.9	5.6
2	14	2	$1/\lambda_i$	93.8	5.0
3	3	1	1	94.3	4.8
3	5	1	1	91.3	5.2
3	14	1	1	92.1	5.4

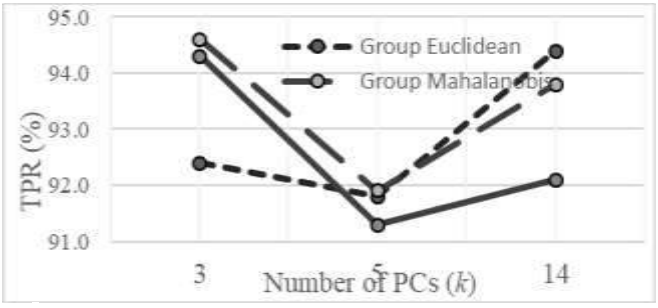
Table III shows the number of anomalous flows that are true detected in dataset 1 for case 1, 2 and 3, respectively. In fact, the total number of anomalous flows in dataset 1 is 84476.

TABLE III
NUMBER OF TRUE DETECTED ANOMALY FLOWS (TP), DATASET 1

Case	Parameters	TP			
		k=3	*k=4*	*k=5*	*k=14*
1	*c=2, w_i=1*	78106	77430	77611	80082
2	*c=2, w_i= 1/λ_i*	79951	78626	77633	79302
3	*c=1, w_i=1*	79783	77678	78465	78431

Figure 3 shows the number of true detected anomalous flows in dataset 1 for three cases regarding different *k* (number of PCs)..

Fig. 3. TP for Anomalous Flows in Dataset 1.



.

Fig. 4. TPR Comparison using Dataset 1.

Figure 4 shows the comparison of TPR between our scheme and typical previous works according to the number of PCs used. As presented in the Section II, the Group Euclidean represents the TPR using Euclidean distance formula, which is used in previous works such as [7, 11, 17, 24-26]. The Group Mahalanobis represents the TPR using Mahalanobis distance formula, which is used in previous works such as [10, 12, 14-16, 18, 20-22].

As showed in figure 4, our scheme provides acceptable TPR (94.3%) in comparison to other works by *k*=3. The Group Mahalanobis gives better TPR (94.6%), but the complexity of our scheme is O(3), while the complexity of Group Mahalanobis is $O(w3^2)$. Thus, our scheme can be used for quick detection using *k=3*.

Table IV and Table V show the experiment results with dataset 2 and dataset 3, respectively. Threshold is *95%*. The results are showed for different combination of the parameters k (number of PCs), *c* and w_i. The grey row in the table indicates that our scheme (*k=3, c=1,* w_i =1) provides acceptable TPR and FPR with lower complexity in comparison to other previous works.

TABLE IV
NUMBER OF TRUE DETECTED ANOMALY FLOWS (TP), DATASET 2

Case	*k*	*c*	w_i	*TPR (%)*	*FPR (%)*
1	3	2	1	95.9	5.4
1	5	2	1	99.9	5.1
1	14	2	1	93.1	4.6
2	3	2	$1/\lambda_i$	93.8	4.7
2	5	2	$1/\lambda_i$	96.9	4.8
2	14	2	$1/\lambda_i$	99.0	4.9
3	3	1	1	95.9	5.0
3	5	1	1	92.3	4.7
3	14	1	1	93.8	5.1

TABLE V
NUMBER OF TRUE DETECTED ANOMALY FLOWS (TP), DATASET 3

Case	*k*	*c*	w_i	TPR (%)	FPR (%)
1	3	2	1	99.4	5.6
1	5	2	1	99.9	5.3
1	14	2	1	99.9	5.4
2	3	2	$1/\lambda_i$	99.7	4.9
2	5	2	$1/\lambda_i$	99.8	5.0
2	14	2	$1/\lambda_i$	99.9	5.3
3	3	1	1	98.6	4.9
3	5	1	1	97.6	5.3
3	14	1	1	100	4.9

Table VI shows the number of normal traffic flows that are true detected in dataset 2 for case 1, case 2 and case 3, respectively. The total number of actual normal flows in dataset 2 is 75358.

TABLE VI
TRUE DETECTED ANOMALY FLOWS (TP) IN DIFFERENCE CASES

Case	Parameters	TN			
		k=3	*k=4*	*k=5*	*k=14*
1	*c=2, w_i=1*	71275	71385	71500	71886
2	*c=2, w_i= 1/λ_i*	71820	71417	71668	71675
3	*c=1, w_i=1*	71586	71699	71767	71504

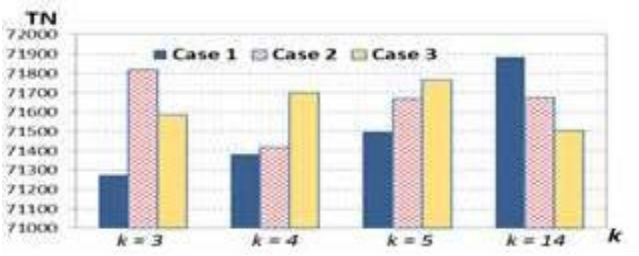


Fig. 5. TN for Normal Flows in Dataset 2.

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Figure 5 shows the number of true detected normal flows in dataset 2 for three cases according to the number of PCs. The reason for TN results is that the last PCs are responsible for the variance of the anomalies.

The experiment results showed that our scheme can provide comparable accuracy with lower complexity (O(kn) instead of O(kn^2) in other schemes) in comparison to typical works using Euclidean distance formula (Group Euclidean) or using Mahalanobis distance formula (Group Mahalanobis). For quick network anomaly detection, we can use the general distance formula with few PCs instead of using all PCs or a large number of PCs as presented in previous works.

V. Conclusion

Network traffic anomaly detection is one of the challenging tasks of network operators, especially in IoT network environment due to many constraints such as limited resource and performance. The detection scheme should be simple (lower complexity) and quick as possible. Although many detection methods have been proposed in the past, methods based on Principal Component Analysis (PCA) received a lot of attention. PCA based methods promise suitable solutions for IoT networks, but the remaining issues need to be investigated including selection of principal components, distance formula, complexity reduction. There is still very few works investigating PCA based methods for network anomaly detection in IoT networks

The paper investigated problems of PCA based methods for IoT networks including: 1) selection of PCs for an effective detection regarding the IoT network constraints, 2) distance measure for lower complexity and 3) quick detection of network traffic anomalies. Based on selecting few PCs using our developed general distance formula, we proposed a novel detection scheme with two levels. The first level is for quick detection with few k principal components in order to reduce the complexity to O(k) while keeping an acceptable detection rate in comparison to previous methods. The second level is for detailed detection with a number of principal components. The paper presented various experiments using three random datasets from Kyoto Honeypot to show the feasibility of our proposed scheme.

Acknowledgment

This work is supported by the ASEAN IVO project “A Hybrid Security Framework for IoT Networks”.

The authors thank the National Institute of Information and Communications (NICT, Japan) and NES (NEC, Japan) for the supports.

References

[1] Q.Jing, et al., "Security of the IoT: Perspectives and Challenges". Wireless Networks, Vol 20, Issue 8, Nov. 2014, pp.2481-2501

[2] Y.M. Pa pa, S. Suzuki, K. Yoshioka, T. Matsumoto, T. Kasama, C. Rossow. "IoTPOT: A Novel Hoeypot for Revealing Current IoT. Threats". Journal of Information Processing, Vol 24, No.3, May 2016, pp.522-533.

[3] A.V. Vijayalakshmi, L. Arockiam, "A Study on Security Issues and Challenges in IoT". Engineering Sciences & Management Research. Vol 3, No 11, Nov. 2016, pp. 34-43

[4] V. Chandola, A. Banerjee, V. Kumar, “Anomaly detection: a survey”, ACM Computing Surveys, 2009, 41, (3), pp. 1-58.

[5] M. Ahmed, A. Mahmood, J. Hu, “A survey of network anomaly detection techniques”, Network & Computer Applications, 2016, 60 (2016), pp. 19-31.

[6] R. Bansal, et al., “Outlier detection: applications and techniques in data mining”, Cloud Syst./Big Data Eng. 2016, pp. 373-377.

[7] D. Weller-Fahy, B. Borgehtti, A. Sodemann, “A survey of distance and similarity measures used within network intrusion anomaly detection”, IEEE Communication Survey & Tutorials, 2015, 17, (1), pp. 70-91.

[8] M. Bhuyan, D. Bhattacharyya, J. Kalita, “Network anomaly detection: methods, systems and tools”, IEEE Com. 2014, 16, 303-336.

[9] M. Molina, et al., “Operational experiences with anomaly detection in backbone networks”, Computer & Security, 2012, 31, (3), pp. 273-285.

[10] F. Harrou, et al.,“Amalgamation of anomaly detection indices for enhanced process monitoring”, Loss Prevention in the Process Industries, 2016, 40, (3), pp. 365-377.

[11] H. Om, T. Hazra, “Statistical techniques in anomaly intrusion detection system”, Advanced in Eng. & Technology, 2012, 5, (1), pp.387-398.

[12] M. Shyu, S. Chen, K. Sarinnapakorn, L. Chang, “A novel anomaly detection scheme based on principle component classifier”, Proc. IEEE foundation and New Directions of Data Mining Workshop, Florida, USA, Nov. 2003, pp. 172-179.

[13] A. Lakhina, M. Crowella, C. Diot, “Diagnosing network-wide traffic anomalies”, Proc. ACM SIGCOMM '04, Portland, Oregon, USA, Aug. 2004, pp. 219-230.

[14] D. Brauckhoff, K. Salamatian, M. May, “Applying PCA for traffic anomaly detection: problems and solutions”, Proc. INFOCOM’09, Rio de Janeiro, Brazil, Apr. 2009, pp. 2866-2870.

[15] H. Ringberg, A. Soule, J. Rexford, C. Diot, “Sensitivity of PCA for traffic anomaly detection”, Proc. ACM SIGMETRICS ’07, San Diego, USA, Jun.2007, pp. 109–120.

[16] J. Camacho, A. Perez-Villegas, P. Garcia-Teodoro, G.Macia-Fernandez, “PCA-based multivariate statistical network monitoring for anomaly detection”, Computer & Security, 2016, 59 (2016), pp. 118-137.

[17] Y. Liu, L. Zhang, Y. Guan, “Sketch-based streaming PCA algorithm for network-wide traffic anomaly detection”, Proc. IEEE 30th Int. Conf. on Dist. Computing Systems, June 2010, Genoa, Italy, pp. 807-816.

[18] C. Callegari, L. Gazzarrini, S. Giordano, M. Pagano, T. Pepe, “A novel PCA-based network anomaly detection”, Prof. IEEE Int. Conf. on Communications, Kyoto, Japan, Jun. 2011, pp. 1-5.

[19] M. Bhuyan, D. Bhattacharyya, J. Kalita, “A multi-step outlier-based anomaly detection approach to network-wide traffic”, Information Sciences 2016, 348 (2016), pp. 243-271.

[20] U. Kwitt, Hofmann, “Robust methods for unsupervised PCA-based anomaly detection”, Proc. IEEE/IST Work-shop on Monitoring, Attack Detection and Mitigation, Tubingen, Germany, 2006, pp. 1-3.

[21] A. Das, S. Misra, S. Joshi, J. Zambreno, G. Memik, A. Choudhary, “An efficient FPGA implementation of principle component analysis based network intrusion detection system”, Proc. Design, Automation and Test in Europe, Munic, Germany, Mar. 2008, pp. 1160-1165.

[22] G. Zargar, T. Baghaie, “Category-based intrusion detection using PCA”, Journal of Information Security, 2012, (3), pp. 259-271.

[23] Y. Lee, Y. Yeh, Y. Wang, “Anomaly detection via online oversampling principal component analysis”, IEEE Trans. On Knowledge & Data Engineering, 2013, 25, (7), pp. 1460-1470.

[24] M. Elrawy, T. Abdelhamid, A. Mohamed, “IDS in telecommunication network using PCA”, International Journal of Computer Networks & Communications, 2013, 5, (4), pp.147-157.

[25] D. Bayarjargal, G. Cho, “Detecting an anomalous traffic attack area based on entropy distribution and Mahalanobis distance”, International Journal of Security and Its Applications, 2014, 8, (2), pp. 87-94.

[26] H. Huang, H. Al-Azzawi, H. Brani, “Network Traffic Anomaly Detection”, arXiv:1402.0856 preprint, 2014.

[27] Z. Fan, Y. Xu, W. Zuo, J. Yang, J. Tang, Z. Lai, D. Zhang, “Modified principal component analysis: an integration of multiple similarity subspace models”, IEEE Trans. On Neural Networks & Learning Systems, 2014, 25, (8), pp. 1538-1552.

[28] H.D. Nguyen, D.H. Hoang, “A model for network traffic anomaly detection”, ICACT Trans. on Advanced Communications Technology, 2015, 4, (4), pp.644-650.

[29] E.G. Nascimento, O. Tavares, A.F. Souza, “A cluster-based algorithm for anomaly detection in time series using Mahalanobis distance”, Proc. Int. Conf. Artificial Intelligence, July 2015, Nevada, USA, pp. 622-628.

[30] S. Franklin, M. Brodeur, “A practical application of a robust multivariate outlier detection method”, Proc. of Survey research methods section, American Statistical Association, 1997, pp. 186-191, http://www.amstat.org/ sections/srms/proceedings.

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

[31] M. Pasha, S.M. Waqas Shah, U. Pasha, "Security Framework for IoT Systems" International Journal of Computer Science and Information Security (IJCSIS), Vol 14, No 11, Nov. 2016, pp.99-104.
[32] A.V. Vijayalakshmi, L. Arockiam. "A Study on Security Issues and Challenges in IoT", Intl. Journal of Engineering Sciences & Management Research. Vol 3, No 11, Nov. 2016, pp. 34-43.
[33] N.C. Winget, Aa.R. Sadeghi, Y. Jin. "INVITED: Can IoT be Secured: Emerging Challenges in Connecting the Unconnected". DAC' 2016, June 05-09, 2016, Austin, TX, USA.
[34] H.H. Pajough, R. Javidan, R. Khayami, D. Ali, K.K. Raymond Choo, "A Two-layer Dimention Reduction and Two-tier Classification Model for Anomaly-based Intrusion Detection in IoT Backbone Networks", IEEE Trans. On Emerging Topics in Computing, 29 Nov. 2016.
[35] S.K. Sharma, X. Wang, "Live Data Analytics with Collaborative Edge and Cloud Processing in Wireless IoT Networks", IEEE Access, Vol 5, March 2017, pp. 4621-4635.
[36] R. Ferrando, P. Stacey, "Classification of Device Behaviour in Internet of Things Infrastructures", Proc. of International Conference on Internet of Things and Machine Learning, Oct. 2017.
[37] S. Zhao, W.Li, T. Zia, A.Y. Zomaya, "A Dimension Reduction Model and Classifier for Anomaly-Based Intrusion Detection in Internet of Things", IEEE 15th International Conference on Pervasive Intelligence and Computing (PICom 2017), Nov. 2017.
[38] J.P. Geer, "Some aspects of Minkowski distance", research report, University of Leiden (1995), pp. 1-31.
[39] H.D. Nguyen, D.H. Hoang. "Network traffic anomaly detection in the condition of noisy training dataset". Journal of Science & Technology on Information and Communications, Vol 1, CS01, 2016, pp.5-18.
[40] Traffic data from Kyoto University's Honeypots, http://www.takakuara.com/Kyoto_data/

Dang Hai Hoang, A/Prof. Dr. DSc., PhD (1999), DSc (2002) at TU Ilmenau, Germany. Current institution: Posts and Telecommunication Institute of Technology. Research interests: Communication network, IoT networks, information security, network security.

Ha Duong Nguyen, BSc (2001), MSc (2003) at TU Hanoi, PhD (2017) at PTIT, Vietnam. Current institution: Faculty of Information Technology. Research interests: Communication network, telecommunication networks, information security, network security.

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

SEFL: Selective Ensemble Fuzzy Learner for Cognitive Detection of Bio-Modality Spoofing in MCPS

Nishat I Mowla*, Inshil Doh**, Kijoon Chae*

*Department of Computer Science and Engineering, Ewha Womans University, 52, Ewhayeodaegil, Seodaemungu, Seoul, 03760, Korea

**Department of Cyber Security,Ewha Womans University, 52, Ewhayeodaegil, Seodaemungu, Seoul, 03760, Korea Name

nishat.i.mowla@gmail.com, isdoh1@ewha.ac.kr, kjchae@ewha.ac.kr

***Abstract*—User authentication in a Medical Cyber Physical Systems (MCPS) can be effectively done using biometric features. Biometric features, widely used for user authentication, are equally important to national and global technology systems. Biometric features, such as face, iris, fingerprint, are commonly used while more recently palm, vein and gait are also getting attention. To fail the traditional biometric detection systems, various spoofing approaches have also been developed over time. Among various methods, image synthesis with play-doh, gelatin, ecoflex etc. are some of the more common ways for spoofing bio-modalities. Success of traditional detection systems are related to custom tailored solutions where feature engineering for each attack type must be developed. However, this is not a feasible process when we consider countless attack possibilities. Also, a slight change in the attack can cause the whole system to be redesigned and therefore becomes a limiting constraint. The recent success of machine learning inspires this paper to explore weak and strong learners with ensemble learning approaches using AdaBoost. In essence, the paper proposes a selective ensemble fuzzy learner approach using Ada Boost, feature selection and combination of weak and strong learners to enhance the detection of bio-modality spoofing for MCPS. Our proposal was experimented on real datasets and verified on the fingerprint and iris benchmark.**

***Keyword*—MCPS, Biometric spoofing, Spoofing Detection, Ensemble Learning, Feature selection**

I. INTRODUCTION

MEDICAL Cyber Physical Systems are a four-layer architecture which extends from users in the acquisition layer to cloudlet to cloud and then to the caregiver. As it is a sensory network composed of medical aspects, biometric features can be effectively used for user authentication. An increasing interest in the evaluation of biometric systems in recent years have been observed for user identification and authentication leveraging traditional fingerprint, iris and more recently vein, blood flow etc. In the other hand, various spoofing attacks have also been developed to defeat these systems. Given these attacks are performed in an analog domain with heavy device dependency following regular protocol, digital protection schemes such as encryption, digital signature or watermarking are unsuitable. The success of detection mechanism for such types of attacks are often linked to custom tailored solution where feature engineering plays a major role. Nevertheless, in a pool of myriad possible attacks, a slight change in the attack model can cause the whole mechanism to be redesigned to adopt the new mechanism which surely becomes a limiting constraint [1]. The past few years have witnessed the success of Deep Learning techniques such as Convolutional Neural Network (CNN) for efficient image classification and Recurrent Neural Network (RNN) for its special recurrence capabilities in contextual image recognition [2]. While high accuracy can be achieved, the computation can take long time which becomes a limiting constraint for time-critical systems such as MCPS.

Previously, we proposed a fuzzy ensemble learner based cognitive detection scheme for fingerprint spoofing in an MCPS where a selective ensemble architecture was proposed using Ada Boost and feature selection [31]. In this paper, we extend the research to more complex modality of the iris benchmark to further verify the effectiveness of the selective ensemble architecture to make the problem space simpler and enhance the overall performance. The performance of the strong learner is further substantiated with the more complex representation of the iris benchmark to provide further insight.

We discuss some of the related works in sections II. In section III we discuss our proposed mechanism and evaluation results followed by the conclusion in section IV.

Manuscript received on Jan. 15, 2018. This work is sponsored by Basic Science Research Program through the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIP), and a follow-up of the invited journal to the accepted & presented paper of the 19th International Conference on Advanced Communication Technology (ICACT2017), and Grant ID is 2016R1A2B4015899. Kijoon Chae is the corresponding author.

Chae. Kijoon. Author is with Ewha Womans University, Seoul, 120750 Korea (corresponding author to provide phone: +82-10-3726-6157; e-mail: kjchae@ ewha.ac.kr).

Mowla Nishat. Author, is with Ewha Womans University, Seoul, 120750 Korea. (e-mail: nishat.i.mowla@gmail.com).

Doh Inshil. Author is with Ewha Womans University, Seoul, 120750 Korea. (e-mail: isdoh1@ewha.ac.kr).

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

II. Related Works

In this section, we review anti-spoofing related work for bio-identifiable modalities focusing on fingerprint spoofing.

A. Bio-identifiable Modality Spoofing Detection

For fingerprint spoofing detection, both hardware-based (exploring extra sensors) and software-based solutions (relying only on the information acquired by the standard acquisition sensor of the authentication system) have been explored. In [3] quality measures such as ridge strength or directionality, ridge continuity, ridge clarity, and integrity of the ridge-valley structure was used as a set of features for fingerprint liveness detection. In [4], the presence of gummy fingers was explored using various methods. In [5], a method for representing all spectrum characteristics in a compact feature representation form was explored. In [6], well suited to high contrast patterns such as the ridges and valleys of fingerprints images, Weber Local Image Descriptor (WLD) for liveness detection was considered. In [7] Multi-Scale Block Local Ternary Patterns (MBLTP) was proposed as a liveness detection scheme. In [8], Binarized Statistical Image Features (BSIF) was explored which was originally proposed in [9]. According to reports in the LivDet 2013 Fingerprint Liveness Detection Competition [28], the fingerprint spoofing attack detection task remains an open problem with results still far from a perfect classification rate. Mostly hard-coded features, sometimes exploring quality metrics related to the modality (e.g., directionality and ridge strength), general texture patterns (e.g., LBP-, MBLTP-, and LPQ-based methods), and filter learning through natural image statistics are heavily followed.

For iris spoofing spoofing detection, the use of Fast Fourier Transform was proposed in [10] to verify the high frequency spectral magnitude in the frequency domain for iris spoofing detection. Solutions for iris liveness detection range from hardware-based solutions [11] [12] [13] to software-based solutions relying on texture analysis for detecting an attacker using contact lenses with someone else's printed pattern [14]. Software-based solutions considering cosmetic contact lenses [15], [16], [17], [18]; pupil constriction [19]; and multi biometrics of electroencephalogram (EEG) and iris together [20] had been explored rigorously. In [21], best features are selected through sequential floating feature selection (SFFS) [22] to feed a quadratic discriminant classifier. In [23], image quality measures were explored, and three classification techniques were proposed. In [24], the previous work was extended by using a feature selection step on the features of the studied methods to obtain the "best features" and then used well-known classifiers for the decision making. In [25], iris segmentation was proposed to obtain the iris contour and feature extraction processes were adapted to the resulting non-circular iris regions. In [26], a general framework for iris image classification based on a Hierarchical Visual Codebook (HVC) encoding the texture primitives of iris images was proposed. For iris spoofing detection, features have been profoundly studied through image-quality metrics, texture patterns, bags-of-visual-words and noise artifacts.

B. Machine Learning based Bio-modality Spoofing Detection

A couple of machine learning based methods was proposed for bio-modality spoofing detection in recent years. A deep architecture based on Convolutional Neural Network using Architecture Optimization (AO) and Filter Optimization (FO) was proposed in [1]. The results were verified in three different modalities and in multiple real datasets. In [2], a combination of Recurrent Neural Network and Convolutional Neural Network was used to create a deep architecture which is context aware as well as capable of detecting high level features. Different existing datasets are used to evaluate the performance bio-identifiable modality spoofing using machine learning algorithms in many research works [1]. However, these architectures require heavy weight algorithms which cannot run fast enough without the use of GPU which incurs another level of computational cost. In [27], a lightweight Adaboost based model with k-means clustering was used to optimize detection accuracy of images. While some considered various deep and boosting algorithms, not many explored the role of features and ensemble learning in a profound way. Our proposed approach, therefore, aims to leverage low computation weak learners and moderately strong learners with boosting ensemble and feature selection for detecting fingerprint and iris benchmark spoofing detection as an authentication verification scheme in Medical Cyber Physical Systems.

III. Proposed Mechanism

Machine Learning performs various levels of learning. It has effectively explored various boosting techniques such as Ada Boost. The main idea of Ada Boost is to set a bunch of weak classifiers or learners working together to outperform a single strong learner [27]. The idea can also be verified with various weak learner architectures such as one-level Decision Tree. Boosting is also a form of ensemble learning where the output of the most efficient learner is incorporated into the final result [27]. However, there are certain features that are learned well by weak learners and certain features that are learned better by strong learners. Based on such facts, our paper is inspired to propose training light weight strong learners parallelly with the weak learners, as a combo learner, which undergoes selective ensemble learning with Ada Boost and feature selection in order to come up with the final decision algorithm. The following sub-section describes our scheme in more detail.

In our proposed scheme, the bio-modality spoofing is learned with one strong learner and one weak learner using Ada Boost. Both learners are represented in different feature selected environments. The fuzzy ensemble algorithm, then, chooses the best combination of the machine learning algorithm and feature selection value as the final decision-making algorithm. Fig 1 shows the basic workflow of our proposed fuzzy ensemble learning algorithm using Ada Boost, feature selection and strong and weak learner combinations.

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

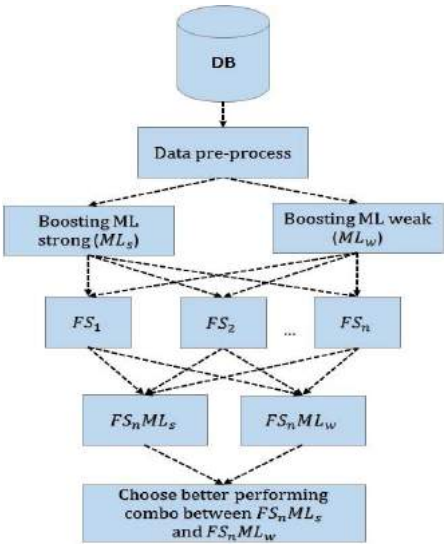


Fig. 1 Selective Ensemble Fuzzy Learning [31].

In the above figure, the DB is used to provide a two-class problem containing spoofed bio-identifiable modality and live bio-identifiable modality. The two-class problem is then learned by one strong learning algorithm and one weak learning algorithm. FS_1, FS_2, FS_n are the different feature selection environments where n is the total number of features used in the experimentation. FS_nML_s represents the best feature selected environment of the strong learning algorithm. Similarly, FS_nML_w represents the best feature selected environment of the weak learning algorithm. Finally, the best combination out of these two environments, FS_nML_s and FS_nML_w, is selected as the decision-making algorithm combination for the final decision process.

The combination of weak and strong learners with Ada Boost is an unexplored field with potential. This is because, it is believed that weak learners together can perform better but it is not explored how a moderately strong learner with lower computational overhead added to this environment could help the overall detection performance. Besides, ensemble learning is computationally expensive. Hence, selective ensemble learning is proposed which can reduce the number of features. The latter is further enhanced by introducing a feature selection methodology to aid in the selective ensemble approach.

LiveDet 2015 [28] and Warsaw [29] dataset provides different kinds of modality spoofing benchmark along with live dataset benchmark. On these different spoofing datasets, we apply our selective fuzzy ensemble learning with feature selection and Ada boost which, as discussed before, can be summarized to follow three major steps as discussed below.

***1) Fuzzy Learning*:** One strong learner and one weak learner is used to learn the spoofed and non-spoofed instances. Both learners are processed with boosting algorithm by using Ada Boost.

2) Feature Selection: This is done in two inter-linked steps. Both learners in the fuzzy learning approach is represented in different feature selected environments. It is known that best performance is not always obtained by using the most number of features but by using the most relevant features. Therefore, the best feature selection value is saved for both the weak learner and the strong learner.

3) *Combo Selection:* The feature selection value with higher accuracy, among the two best feature selection values saved in the feature selection step, is selected. Therefore, the final decision-making algorithm uses this feature selection value along with its corresponding machine learning algorithm which could either be a weak or a strong learning algorithm based on the attained accuracy.

The algorithm of our proposed selective fuzzy ensemble learner using Ada Boost and feature selection is shown below.

Algorithm 1 Selective Ensemble Fuzzy Boosted Learner

```
pre-process images to form a 2-class problem;
initialize class spoofed = 1;
initialize class live =0;
run Feature Selection algorithm;
    run classification with boosted strong learner;
    get best learner and feature selection combo, fs_x1;
    run classification with boosted weak learner;
    get best learner and features selection combo, fs_x2;
    if fs_x1 > fs_x2 then
        choose fs_x1;
    else
        choose fs_x2;
```

After pre-processing the data to form a 2-class problem where one is the spoofed class and the other is the non-spoofed class. Class spoofed is initialized as 1 and class live or non-spoofed is initialized as 0. The feature selection algorithm is run for both the strong and weak learner. The combo of the learner and feature selection is compared to find the best combo.

IV. Peroformance Evaluation

The LiveDet 2015 [28] and Warsaw dataset [29] dataset provides a set of spoofed and live biometric dataset. For our experimentation, we used the benchmark for the fingerprint and iris dataset containing spoofed and live instances. The benchmarks were pre-processed to extract features from each spoofed and live image and stored as ARFF (Attribute-Relation File Format) files.

After pre-processing the benchmarks to extract the features, a two-class problem is created containing spoofed and non-spoofed instances. We used Knime Analytics tool[30] for pre-processing the dataset and simulating the boosted strong and weak learners. Therefore, the dataset is then learned by one boosted strong learner and one boosted weak learner. For strong learner we used Naïve Bayes and for weak learner we used a One-Level Decision Tree. Fig. 4 and Fig. 5 shows our simulation of the dataset classification with the learners using Ada Boosting and feature selection.

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

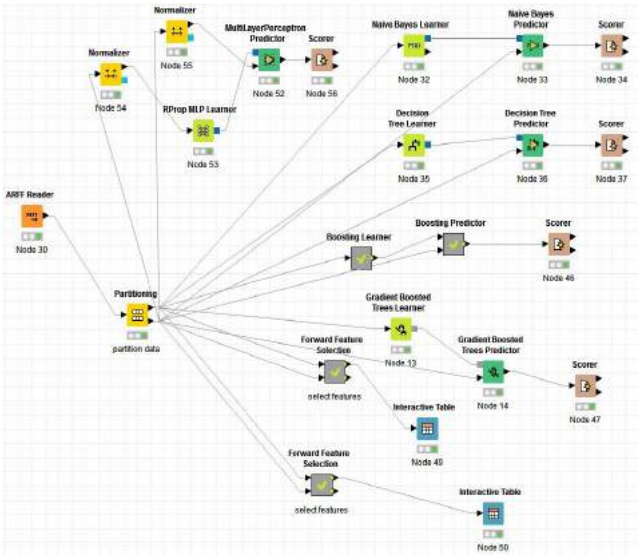


Fig. 2. Data partitioning and forwarding feature selection [31].

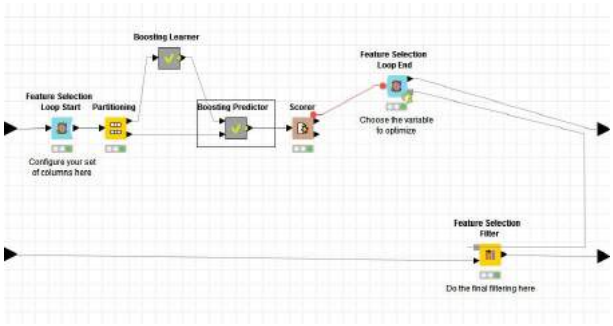


Fig. 3. Boosting Learner and Feature Selction [31].

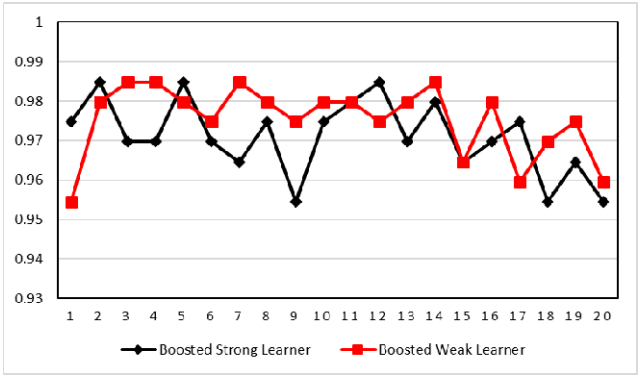


Fig. 4. Different feature selected values for Boosted Strong and Weak Learner for fingerprint benchmark.

The learning is performed for a range of feature selected environments ranging from 1-feature selection to all 20-feature selection. Fig. 4 and Fig. 5 shows the results for the selective ensemble strong boosted learner, Naïve Bayes and weak boosted learner, 1-level Boosted Decision Tree, in different feature selected environments for fingerprint and iris benchmark respectively.

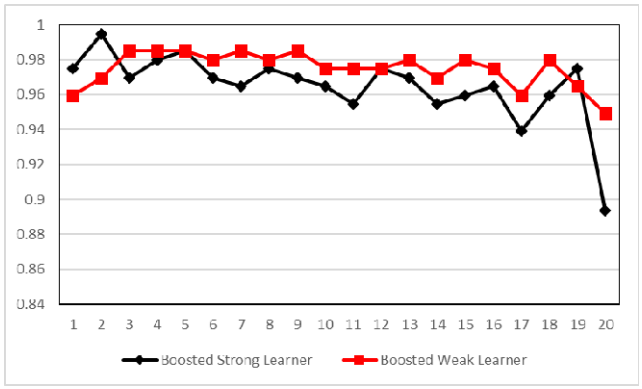


Fig. 5. Different feature selected values for Boosted Strong and Weak Learner for iris benchmark.

The combined learning of these two learners are monitored for all feature selection environment to extract the best selection values. As can be seen from Fig. 4, in terms of average accuracy, we get highest performance (98.485%) with 2,5 and 12 feature selection value for the boosted strong learner and 3,4,7 and 14 feature selection value for the boosted weak learner. Combinedly, we can get achieve highest performance with 2,3,4,5,7,12 and 14 features. From Fig. 5, we get highest performance (99.495%) with 2 features for boosted strong learner and with 3,4,5,7 and 9 features for the boosted weak learner. Thus, based on the available features and preferred algorithm between the strong and weak learner, we can choose the final combination as the decision algorithm. Table 1 summarizes a list of comparable algorithms to our proposed Selective Ensemble Fuzzy Boosted Learner algorithm.

TABLE I
BIO-MODALITY DETECTION OF FINGERPRINT BENCHMARK

Algorithm	Average Accuracy	Number of Features
Naïve Bayes	92.407	20
Decision Tree	96.418	20
Boosted Naïve Bayes	94.628	20
Boosted Decision Tree	97.278	20
Multi-Layer Perceptron	95.989	20
Selective Fuzzy	**98.485**	Min: 2
Ensemble Learner		Max: 14

TABLE II
BIO-MODALITY DETECTION OF IRIS BENCHMARK

Algorithm	Average Accuracy	Number of Features
Naïve Bayes	92.407	20
Decision Tree	96.418	20
Boosted Naïve Bayes	89.393	20
Boosted Decision Tree	94.949	20
Multi-Layer Perceptron	81.5	20
Selective Fuzzy	**99.495**	Min: 2
Ensemble Learner		Max: 2

As can be seen from the above two tables, the learning is optimum for our proposed selective ensemble fuzzy learner in terms of average accuracy metric. For fingerprint, this high performance can be achieved by the lowest feature selection using 2 features with a strong boosted learner, or the highest feature selection using 14 features with a weak boosted learner.

In this case, the highest number of features used is 14 which is still lower than the total number of features that is used by the

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

other algorithms. Besides 5 features and 12 features with a strong boosted learner and 3,4 and 7 features with a weak boosted learner can also be used giving the same highest accuracy of 98.485 percent. In our approach, since the number of features used will be less than the total number of features and the algorithm can alternate between a weak and moderately strong learner, the overall overhead can be significantly reduced in best cases with 2,3,4 and 5 features, moderate cases with 7 features and in worst cases with 12 or 14 features which is still lesser than the total number of features. For iris, the highest performance can be achieved by the feature selection using 2 features with a strong boosted learner. Since this feature selection alone provides the highest performance, we do not need to consider other feature selection and learner combos. As can be seen, in this case, a strong boosted learner outperforms a weak boosted learner. And since we can alternate between the strong learner and weak learner we can utilize both the learners depending on the kind of learning environment.

The performance gain of this paper is credited to the fact that learning is flexible and able to move between a moderately strong and a weak learner. While the statement, weak learners together can perform better than a strong learner holds, it is also shown that strong learners working together with weak learners can also gain promising performance in desired feature selected values. Ensemble learning is computationally expensive but, in our case, the feature selected ensemble is helping to make the problem space smaller and simpler. Thus, while the performance is improving, the overall complexity is not rising significantly.

V. Conclusion

In this paper, we have proposed a selective ensemble fuzzy learner with Ada Boost and Feature Selection in order to detect bio-identifiable modality spoofing of fingerprint and iris benchmark that can be used for authentication in a Medical Cyber Physical System. We have shown that our proposed mechanism enhances the performance of the traditional boosted learning algorithms. Our mechanism also considers a selective ensemble learning approach to reduce the overall computational overhead. In future work, we hope to apply our proposed mechanism in other bio-identifiable modality spoofing.

Acknowledgment

The work was supported by the National Research Foundation of Korea (NRF) funded by the Korea government (MSIP) (2016R1A2B4015899). Kijoon Chae is corresponding author

References

[1] D. Menotti, G. Chiachia, A. Pinto, W. R. Schwartz, H. P. A. X. Falcao, and A. Rocha. "Deep representations for iris, face, and fingerprint spoofing detection." *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 864-879, 2015.

[2] M. Liang, and X. Hu. "Recurrent convolutional neural network for object recognition." *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3367-3375. 2015.

[3] J. Galbally, F. Alonso-Fernandez, J. Fierrez, and J. Ortega-Garcia, "Fingerprint liveness detection based on quality measures," *in Proc. Int. Conf. Biometrics, Identity, Secur. (BIdS)*, pp. 1–8, 2009.

[4] J. Galbally, F. Alonso-Fernandez, J. Fierrez, and J. Ortega-Garcia, "A high performance fingerprint liveness detection method based on quality related features," *Future Generat. Comput. Syst.*, vol. 28, no. 1, pp. 311–321, 2012.

[5] L. Ghiani, G. L. Marcialis, and F. Roli, "Fingerprint liveness detection by local phase quantization," *in Proc. 21st Int. Conf. Pattern Recognit. (ICPR)*, Nov. 2012, pp. 537–540.

[6] D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva, "Fingerprint liveness detection based on weber local image descriptor," *in Proc. IEEE Workshop Biometric Meas. Syst. Secur. Med. Appl.*, pp. 46–50, Sep. 2013.

[7] X. Jia et al., "Multi-scale block local ternary patterns for fingerprints vitality detection," *in Proc. IAPR Int. Conf. Biometrics (ICB)*, pp. 1–6, 2013.

[8] L. Ghiani, A. Hadid, G. L. Marcialis, and F. Roli, "Fingerprint liveness detection using binarized statistical image features," *in Proc. IEEE Int. Conf. Biometrics, Theory, Appl., Syst. (BTAS)*, pp. 1–6, Sep./Oct. 2013.

[9] J. Kannala and E. Rahtu, "BSIF: Binarized statistical image features," *in Proc. 21st Int. Conf. Pattern Recognit. (ICPR)*, pp. 1363–1366, Nov. 2012.

[10] J. Daugman, "Recognizing persons by their iris patterns," *in Biometrics:Personal Identification in Networked Society*, Boston, MA, USA:Kluwer, pp. 103-121, 1999.

[11] C. Lee, K. R. Park, and J. Kim, "Fake iris detection by using purkinje image," *in Advances in Biometrics (Lecture Notes in Computer Science), New York, NY, USA: Springer-Verlag*, vol. 3832, pp. 397-403, 2005.

[12] A. Pacut and A. Czajka, "Aliveness detection for iris biometrics," *in Proc. 40th Annu. IEEE Int. Carnahan Conf. Secur. Technol.*, pp. 122-129, 2006.

[13] M. Kanematsu, H. Takano, and K. Nakamura, "Highly reliable liveness detection method for iris recognition," *in Proc. Annu. Conf. SICE*, pp. 361-364, 2007.

[14] Z. Wei, X. Qiu, Z. Sun, and T. Tan, "Counterfeit iris detection based on texture analysis," *in Proc. 19th Int. Conf. Pattern Recognit. (ICPR)*, pp. 1-4, 2008.

[15] K. W. Bowyer and J. S. Doyle, "Cosmetic contact lenses and iris recognition spoofing," *in Computer*, vol. 47, no. 5, pp. 96-98, 2014.

[16] D. Yadav, N. Kohli, J. S. Doyle, R. Singh, M. Vatsa, and K. W. Bowyer, "Unraveling the effect of textured contact lenses on iris recognition," *in IEEE Transactions on Information Forensics and Security*, vol. 9, no. 5, pp. 851-862, 2014.

[17] N. Kohli, D. Yadav, M. Vatsa, and R. Singh, "Revisiting iris recognition with color cosmetic contact lenses," *in Proc. IAPR Int. Conf. Biometrics (ICB)*, pp. 1-7, 2013.

[18] J. S. Doyle, K. W. Bowyer, and P. J. Flynn, "Variation in accuracy of textured contact lens detection based on sensor and lens pattern,"*in Proc. IEEE 6th Int. Conf. Biometrics, Theory, Appl., Syst. (BTAS)*, pp. 1-7, 2013.

[19] X. Huang, C. Ti, Q.-Z. Hou, A. Tokuta, and R. Yang, "An experimental study of pupil constriction for liveness detection," *in Proc. IEEE Workshop Appl. Comput. Vis. (WACV)*, pp. 252-258, 2013.

[20] T. Kathikeyan and B. Sabarigiri, "Countermeasures against IRIS spoofing and liveness detection using Electroencephalogram (EEG)," *in Proc. Int. Conf. Comput., Commun., Appl. (ICCA)*, pp. 1-5, 2012.

[21] J. Galbally, J. Ortiz-Lopez, J. Fierrez, and J. Ortega-Garcia, "Iris liveness detection based on quality related features," *in Proc. IAPR Int. Conf. Biometrics (ICB)*, pp. 271-276, 2012.

[22] P. Pudil, J. Novovicova, and J. Kittler, "Floating search methods in feature selection," *in Pattern Recognit. Lett.*, vol. 15, no. 11, pp. 1119-1125, 1994.

[23] A. F. Sequeira, J. Murari, and J. S. Cardoso, "Iris liveness detection methods in mobile applications," *in Proc. Int. Conf. Comput. Vis. Theory Appl. (VISAPP)*, pp. 22-33, 2014.

[24] A. F. Sequeira, J. Murari, and J. S. Cardoso, "Iris liveness detection methods in the mobile biometrics scenario," *in Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, pp. 3002-3008, 2014.

[25] J. C. Monteiro, A. F. Sequeira, H. P. Oliveira, and J. S. Cardoso, "Robust iris localisation in challenging scenarios," *in Computer Vision, Imaging and Computer Graphics: Theory and Applications (Communications in Computer and Information Science), Berlin, Germany: Springer-Verlag*, 2004.

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

[26] Z. Sun, H. Zhang, T. Tan, and J. Wang, "Iris image classification based on hierarchical visual codebook," *in IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 6, pp. 1120-1133, 2014.
[27] F. Smeraldi, M. Bicego, M. Cristani, and V. Murino. "CLOOSTING: CLustering Data with bOOSTING." *In MCS*, pp. 289-298. 2011.
[28] L. Ghiani et al., "LivDet 2013—Fingerprint liveness detection competition," *in Proc. Int. Conf. Biometrics (ICB)*, pp. 1–6, 2013. [Online]. Available: http://prag.diee.unica.it/fldc/
[29] A. Czajka, "Database of iris printouts and its application: Development of liveness detection method for iris recognition," *in Proc. 18th Int. Conf. Methods Models Autom. Robot. (MMAR)*, pp. 28-33, 2013.
[30] KNIME, Knime Analytics Platform. Available at https://www.knime.com/knime-analytics-platform
[31] N. Mowla, I. Doh, K. Chae, "Selective fuzzy ensemble learner for cognitive detection of bio-identifiable modality spoofing in MCPS", *in 20th International Conference on Advanced Communication Technology (ICACT)*, pp. 63-37, 2018.

Nishat Mowla was born on 1st August, 1989. She received the B.S degree in computer science from Asian University for Women, Chittagong, Bangladesh in 2013, an M.S. degree in computer science and engineering from Ewha Womans University, Seoul, Korea in 2016.

She worked at Asian University for Women, Chittagong, Bangladesh as a Senior Teaching Fellow. She is currently a Ph.D. student at Ewha Womans University, Seoul, Korea. Her research interests include next generation network security, IoT network security and network traffic analysis.

Ms. Mowla received the best thesis award for her Master's thesis in 2016. She was awarded the best paper award in the Qualcomm 2017 paper competition. She also received the outstanding paper award in the 19th International Conference on Advanced Communication Technology (ICACT) in 2017.

Inshil Doh was born on 3rd March, 1970. She received the B.S. and M.S. degrees in computer science and engineering at Ewha Womans University, Korea, in 1993 and 1995, respectively. She received the Ph.D. degree in computer science and engineering from Ewha Womans University in 2007.

She was a research professor of Ewha Womans University in 2009~2010 and of Sungkyunkwan University in 2011. She is currently an assistant professor of Computer Science and Engineering at Ewha Womans University, Seoul, Korea. Her research interests include wireless network, sensor network security, and M2M network security.

From 1995-1998, Prof. Doh worked in Samsung SDS of Korea to develop a marketing system. Prof. Doh received best paper award in Korea information Processing Society in 2009. Prof. Doh also received best paper award in Korea Institute of Information and Communication Engineering Conference in 2015.

Prof. Chae was born on 22nd October, 1957. He received the B.S. degree in mathematics from Yonsei University in 1982, an M.S. degree in computer science from Syracuse University in 1984. He received a Ph.D. degree in electrical and computer engineering from North Carolina State University in 1990.

He is currently a professor in the Department of Computer Science and Engineering at Ewha Womans University, Seoul, Korea. His research interests include sensor network, smart grid, CDN, SDN and IoT, network protocol design and performance evaluation.

Prof. Chae was the advisory board member of ACM Transactions on Internet Technology from 2000 to 2004. He was also a member of the International Who's Who from 2001 to 2010.

Copyright ⓒ 2018 GiRI (Global IT Research Institute)

Volume 7 Issue 5, September. 2018, ISSN: 2288-0003

ICACT-TACT JOURNAL

GIRI
Global IT Research Institute

1713 Obelisk, 216 Seohyunno, Bundang-gu, Sungnam Kyunggi-do, Republic of Korea 13591
Business Licence Number : 220-82-07506, Contact: tact@icact.org Tel: +82-70-4146-4991