`

Computer Science and Information Technology

`

David C. Wyld,
Dhinaharan Nagamalai (Eds)

Computer Science & Information Technology

- 7th International Conference on Computer Science and Information Technology (CSTY 2021), December 18~19, 2021, Dubai, UAE
- 7th International Conference on Signal and Image Processing (SIGI 2021
- 7th International Conference of Managing Value and Supply Chains (MaVaS 2021)

Published By

AIRCC Publishing Corporation

`

Volume Editors

David C. Wyld,
Southeastern Louisiana University, USA
E-mail: David.Wyld@selu.edu

Dhinaharan Nagamalai (Eds),
Wireilla Net Solutions, Australia
E-mail: dhinthia@yahoo.com

ISSN: 2231 - 5403
ISBN: 978-1-925953-56-5
DOI: 10.5121/csit.2021.112201- 10.5121/csit.2021.112209

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

`

Preface

7th International Conference on Computer Science and Information Technology (CSTY 2021), December 18~19, 2021, Dubai, UAE, 7th International Conference on Signal and Image Processing (SIGI 2021) and 7th International Conference of Managing Value and Supply Chains (MaVaS 2021) was collocated with 7th International Conference on Computer Science and Information Technology (CSTY 2021). The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The CSTY 2021, SIGI 2021 and MaVaS 2021 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically.

In closing, CSTY 2021, SIGI 2021 and MaVaS 2021 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the CSTY 2021, SIGI 2021 and MaVaS 2021.

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

David C. Wyld,
Dhinaharan Nagamalai (Eds)

`

General Chair

Organization

David C. Wyld, Southeastern Louisiana University, USA
Dhinaharan Nagamalai (Eds) Wireilla Net Solutions, Australia

Program Committee Members

Abdel-Badeeh M. Salem, Ain Shams University, Egypt
Abdellatif I. Moustafa, Umm AL-Qura University, Saudi Arabia
Abdelmounaim Aggour, Hassan II University - Casablanca, Morocco
Abderrahim Siam, University of Khenchela, Algeria
Abderrahmane Ez-Zahout, Mohammed V University, Morocco
Addisson Salazar, Universitat Politecnica de Valencia, Spain
Ahmad A. Saifan, Yarmouk University, Jordan
Ahmed Alsabbagh, University of Babylon, Iraq
Ahmed Y. Nada, Al - Quds University, Palestine
Ajayeb Salama Abu-Daabes, Emirates Collage of Technology, UAE
Ajit Singh, Patna University, India
Alexander Gelbukh, Instituto Politecnico Nacional, Mexico
Ali A. Amer, Taiz University, Yemen
Ali Abdrhman Mohammed Ukasha, Sebha University, Libya
Amari Farouk, University of Sfax, Tunisia
Amari Houda, Networking & Telecom Engineering, Tunisia
Amel Ourici, Badji Mokhtar University of Annaba, Algeria
Amizah Malip, University of Malaya, Malaysia
Anand Nayyar, Duy Tan University, Vietnam
Apurv Asthana, University of Michigan, USA
Aridj Mohamed, Hassiba Benbouali University, Chlef Algeria
Ashish Deshmukh, SVKM's, India
Assem abdel hamied moussa, Chief Eng egyptair, Egypt
Atul Garg, Chitkara University, India
Avinash Waikar, Southeastern Louisiana University, USA
Azah Kamilah Muda, UTeM, Malaysia
Basant Verma, Chandigarh Engineering College, India
Batuhan Ayhan, Marmara University, Turkey
Beshair Alsiddiq, Prince Sultan University, Saudi Arabia
Binod Kumar Pattanayak, Siksha O Anusandhan University, India
Bouchra Marzak, Hassan II University, Morocco
Charles Kalinzi, Kyambogo University, Uganda
Christian Mancas, Ovidius University, Romania
Daniela Lopez De Luise, Universidad de Palermo, Argentina
Deepak Kumar, University of Delhi, India
Denivaldo Lopes, Federal University of Maranhao, Brazil
Dokuz, Eylül University, Turkey
Ekbal Rashid, RTC Institute of Technology, India
Ela Goyal, SIES College of Management Studies, India
Elboukhari Mohamed, University Mohamed First, Morocco
F. Abbasi, Islamic Azad University, Iran
Faisal AbuRub, University of Petra, Jordan
Faouzi Kamoun, Zayed University, UAE
Fariba Azizzadeh, Islamic Azad University, Iran

`

`

`

Technically Sponsored by

Computer Science & Information Technology Community (CSITC)

Artificial Intelligence Community (AIC)

Soft Computing Community (SCC)

Digital Signal & Image Processing Community (DSIPC)

`

7th International Conference on Computer Science and Information Technology (CSTY 2021)

7th International Conference on Signal and Image Processing (SIGI 2021)

7th International Conference of Managing Value and Supply Chains (MaVaS 2021)

ANCHOR DENSITY MINIMIZATION FOR LOCALIZATION IN WIRELESS SENSOR NETWORK (WSN)

Nour Zaarour, Nadir Hakem and NahiKandil

Engineering School, UQAT-LRTCS, Rouyn-Noranda, Canada

ABSTRACT

In wireless sensor networks (WSN) high-accuracy localization is crucial for both of WNS management and many other numerous location-based applications. Only a subset of nodes in a WSN is deployed as anchor nodes with their locations a priori known to localize unknown sensor nodes. The accuracy of the estimated position depends on the number of anchor nodes. Obviously, increasing the number or ratio of anchors will undoubtedly increase the localization accuracy. However, it severely constrains the flexibility of WSN deployment while impacting costs and energy. This paper aims to drastically reduce anchor number or ratio of anchor in WSN deployment and ensures a good trade-off for localization accuracy. Hence, this work presents an approach to decrease the number of anchor nodes without compromising localization accuracy. Assuming a random string WSN topology, the results in terms of anchor rates and localization accuracy are presented and show significant reduction in anchor deployment rates from 32% *to* 2%.

KEYWORDS

Wireless sensor network (WSN), anchors, received signal strength (RSS), localization, path-loss exponent (PLE), connectivity.

1. INTRODUCTION

Localization in wireless sensor networks (WSN) is anessential and critical issue. Most WSN applications necessitate the location of the sensor nodes such as in environment surveillance, object tracking, emergency services, asset management, location-based recommendations, and geosocial networks [1] [2]. Knowing the location is not only necessary to identify the geographic origin of events, for example, the location of a fire or the location of the enemy on a battlefield for the deployment of troops, but it can help in various functionalities system, such as geographic routing, network coverage, perimeter search, topology control, and location-based information polling. Moreover, the availability of cheap wireless networks and the surge in adoption of smartphones make the location-based services (LBS) omnipresent. Indoor LBSs promise enormous potential for research organizations to adapt to different indoor applications such as emergency services and assisted health care systems [2].

One of the simplest techniques is to locate the nodes manually when they are deployed in the environment. However, manual localization is costly in time, due to the large number of nodes to be located. Another technique is to use the Global Positioning System (GPS) which provides highly accurate location information, but it may not be feasible for most WSN deployments such as indoor environment deployment [3].

David C. Wyld et al. (Eds): CSTY, SIGI, MaVaS - 2021
pp. 01-11, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.112201

Since sensor nodes are energy constrained, solutions like GPS are not recommended, GPS components available for WSNs are very costly, exceeding almost three times the cost of a sensor node [4]. Likewise, in some hostile or indoor environments, GPS performance will deteriorate significantly and therefore will be unreliable for location [5]. Hence, various techniques and localization algorithms have been proposed in the literature to localize sensors in WSN [6] [7] [8], however, to achieve high accuracy these techniques, a high percentage of anchors whose location relative to a global reference axis are known a priori has to be used. Nevertheless, not every node of deployed WSN can be equipped with localization components, and that due to cost and power consumption reasons. In this work, we present a localization approach aiming to decrease anchor density, hence network cost, while maintaining a high localization accuracy. This is achieved by using the joint parameter and distance estimation approach based on connectivity and received signal strength.

The remainder of this paper is organized as follows: related works and anchor density impact are discussed in section 2. In section 3, the joint parameter and distance estimation approach based on connectivity- RSS is summarized. Localization and simulations results are analyzed in section 4, and a conclusion and perspectives are drawn in section 5.

2. RELATED WORKS

2.1. Basic Localization Methods

Localization techniques are classified in different categories. Figure 1 presents the taxonomy of localization techniques. An anchor-based localization algorithm uses one or more anchors. These nodes provide location information, in the form of beacon messages, to other nodes whose position is unknown so that they can be located, forming a global coordinate system where the location of each node is estimated, hence the localization is absolute. However, in an anchor free technique, the sensors cooperate with their neighbors, without the use of anchors, and form a local coordinate system where the location of each node is estimated, hence a relative localization is [7].

In centralized techniques, anchors collect the measurements of the unknown node to localize and then send them to a central processor to calculate the position of the unknown node. Usually, this type is not very scalable, as the aggregation of required information such as anchor locations and metrics can require many node collaborations, causing unnecessary overhead and even congestion. While, in distributed techniques the target node can only infer its own location based on information collected locally, and independently.

Range-based localization technique uses the measured distance/angle between nodes to estimate the location. Common measurements used for localizing nodes in WSNs are the RSS [7], time of arrival (ToA) [8], time difference of arrival (TDoA) [9], angle of arrival (AoA) [10]. However, range-free localization technique uses the connectivity or pattern matching method to estimate the location. Such as the approximate point-in-triangulation test (APIT) algorithm [12], the distance vector-hop (DV-Hop) algorithm [13], the centroid localization algorithm [11]. The advantage of using range-based techniques is that they have a high accuracy range compared to range-free techniques. However, these techniques are limited because they require additional hardware, which is expensive for large systems. While in range-free techniques, it is not necessary to determine distances directly; instead, they use radio connectivity to calculate the number of hops between nodes and estimate the location using geometry methods. Certain advantages can be obtained by using these techniques which do not require special hardware support; generally, they are cost effective, mainly to the detriment of the level of precision [14].

The fingerprinting technique or scene analysis is another branch of localization technique. It uses the signatures, and is based on a study campaign conducted in the environment where the location system works. In this method the signal characteristics obtained from a set of locations are catalogued in a first phase, called off-line phase, aiming to build the signature database. Several types of signatures [23] can be used: the powers, angles of arrival, arrival time, broadband parameters such effective delay spread or the number of reflected paths of signals received from the fixed base stations. In the second phase called the real time phase, the locations of the node are estimated by comparing the nodes current signal characteristics with those catalogued previously. However, the requirement for generating a signal signature database makes this technique a laborious collection of data during scene analysis or even unachievable for the most scenarios of the WSNs especially in complex environments.

RSS-based methods are ideal for low-cost and low complexity networks, since no additional hardware is needed. However, the exact knowledge of the propagation model is of greatest importance for RSS-based localization or ranging. A previous work presented a hybrid approach which uses the information of the range-free technique (connectivity information) in order to rectify the errors obtained by the range-based method and that by estimating the parameters of the propagation model to better map RSS measurements into inter-node distance estimation [15] however, the latter work does not shed light on keeping high localization accuracy with a low anchor density.

The accuracy of localization technique is greatly affected by the number of anchors and their placement, playing an essential role in the cost of the network. Many studies have investigated optimal number and placement of anchors to increase the localization accuracy [16] [17] [18]. Moreover, they study optimal anchor placement in area-based localization algorithms with the goal of providing the best placement that maximises accuracy. However, to the best of our knowledge, no work has a goal to decrease the number of anchor while keeping high localization accuracy. Hence, the aim of this paper is to present a localization approach with a low anchor density and a high localization accuracy.

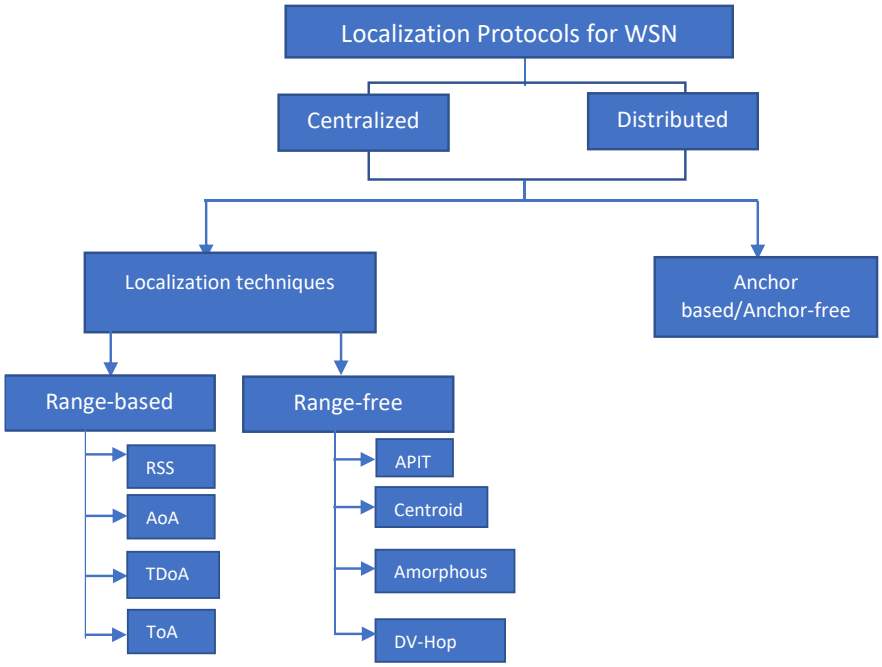

Figure 1. Taxonomy of localization techniques

2.2. Sensor node location

Given a distance measurement between a sensor node and an anchor node, the position of unknown node must be along the circumference of a circle (in two-dimensional space) or sphere (in three-dimensional space) centered at the reference node, with the radius representing the distance between the reference node and the sensor node. At least two reference nodes in one dimension, three non-collinear reference nodes in two dimensions, or four non-coplanar reference nodes in three dimensions are required to obtain a unique location. This process is called trilateration, which assumes perfect distance measurements, which is not achievable in WSNs because of ranging errors.

2.3. Anchor Density Effect

The accuracy of a localization method is governed not only by the efficiency of distance estimation between unknown and anchor node, but also by the number (accuracy increases with anchor percentage) and the position of the anchors themselves [19]. Researchers working on anchor based WSN localization have always been interested in the effect of number and placement of anchor nodes in the network [20]. They have focused on reducing the position error introduced by placement and percentage of anchor nodes in the network. Localization error decreases with the increasing of connected anchor nodes. However, increasing anchor nodes will increase the cost of the deployed network. The main aim in this work is to reduce anchor nodes' percentage while assuring a high localization accuracy.

3. CONNECTIVITY-BASED JOINT PARAMETER ESTIMATION

This section summarizes the approach used to reduce anchor density which proposes a joint estimation scheme for the range, path-loss exponent (PLE), and inter-node distances based on the received signal strength (RSS) and the network's information [15].

3.1. Assumptions

Consider a homogeneous Poisson point process (PPP) in a one-dimensional WSN consisting of N nodes placed randomly at positions x_i for $i = 1, \dots, N$ along the deployment segment $[x_{min}\ x_{max}]$, with node density $\lambda = \frac{N}{(x_{max} - x_{min})}$, and having transmission ranges R_i for $i = 1, \dots, N$. This topology is well-justified in environments that impose one-dimensional deployments such as narrow-vein underground mines [21], sewage or water distribution networks, etc. The received signal power in dBm is modeled as the sum of large-scale path-loss and log- Normal shadowing. The received powerPr_{ij}at node i of a signal emitted from node j is modeled by [22] as determined in equation 1.

$$Pr_{ij}(dij) = P\,r(d_0) - 10\gamma log\left(\frac{d_{ij}}{d_0}\right) + X_\sigma \quad (1)$$

Where $P\,r(d_0)$ is the received power from any given node at the reference distance $d_0 = 1$, γ is the PLE with common values ranging between 2 and 6, $d_{ij}$is the distance separating the two nodes i and j, and X_σ is the large-scale log-Normal shadowing with variance σ^2.

3.2. Poisson Point Process (PPP)

A uniform (homogeneous) PPP is defined in [24] as:

"Let Λ be a locally finite measure on some metric space E. A point processes Φ is Poisson on E if
• For all disjoint subsets $A_1,\cdots,A_n$ of E, the random variables $\Phi(A_i)$ are independent
• For all sets A of E, the random variables $\Phi(A)$ are Poisson"

If a Poisson point process has a constant parameter, λ, then it is considered a homogeneous or stationary PPP [25]. In fact, the parameter λ can be interpreted as the average number of points per unit of length, area or volume, so it is sometimes referred to as the average density.

If two real numbers a and b, such as $a \leq b$, representing points in time, belong to a PPP with parameter $\lambda > 0$, then the probability of n points existing in the interval $(a, b]$ is given by equation 2.

$$P\{N(a,b] = n\} = \frac{[\lambda(b-a)]^n}{n!} e^{-\lambda(b-a)} \tag{2}$$

3.3. Connectivity Information

Two nodes are neighbors at one hop if they are connected, hence, C_{ij} is a random variable presenting the connectivity information defined as in equation 3.

$$C_{ij} = \begin{cases} 1 \; if \; Pr_{ij} \geq P_{th} \\ 0 \; if \; Pr_{ij} < P_{th} \end{cases} \tag{3}$$

Where P_{th} is the power detection threshold.

3.4. Proposed Estimation

3.4.1. PLE estimation

The estimated PLE, $\hat{\bar{\gamma}}$, over the entire wireless sensor network will be estimated by equation 4.

$$\hat{\bar{\gamma}} = \frac{1}{N} \sum_{i=1}^{N} \hat{\gamma}_i \tag{4}$$

Where, $\hat{\gamma}_i$, $\hat{R}_i$, $\hat{\bar{R}}$ are determined in equations 5, 6, and 7 respectively.

$$\hat{\gamma}_i = \frac{-P_{th} + P_r(d_0)}{10 \; log_{10}(\hat{R}_i)} \tag{5}$$

$$\hat{R}_i = \frac{1}{2\lambda} \sum_{j=1}^{N} C_{ij} \tag{6}$$

And

$$\hat{\bar{R}} = \frac{1}{N} \sum_{i=1}^{N} \hat{R}_i \tag{7}$$

3.4.2. Distance estimation

Each node i, for $i = 1, \dots, N$, estimates its distances to its connected neighbor nodes $k \neq i$ as in equation (8).

$$\hat{d}_{ik} = 10^{\frac{Pr(d_0) - Pr_{ik}}{10\,\hat{\gamma}}} \tag{8}$$

Where $\mathrm{Pr}_{\mathrm{ik}}$ is the received power at node i from node k.

4. LOCALIZATION AND SIMULATIONS RESULTS

4.1. Assumptions and WSN Model

The approach presented in this paper consists in decreasing the numberof anchor nodes. To prove its efficiency, a multi-hop linear WSN of N nodes is considered, it is deployed in a homogeneous environment, i.e., all nodes have a priori the same communication range $R_i = R$ for $i = 1, \dots, N$ with density λ. Nodes are positioned in linear topology, on a distance $d = x_{max} - x_{min}$ as shown in figure 2. However, possible extensions to *2D* or *3D* network topologies, beyond the scope of this contribution, are currently under investigation and will be addressed in future publications. Its normalized error (NE), ε_x, is assessed as computed in equation 9.

$$\varepsilon_x = \frac{|(x_i - \hat{x}_i)|}{x_i} \tag{9}$$

Where x_i is the position in one dimension of a node $i, i = 1, \dots, N$, and $\hat{x}_i$ is its estimated position.

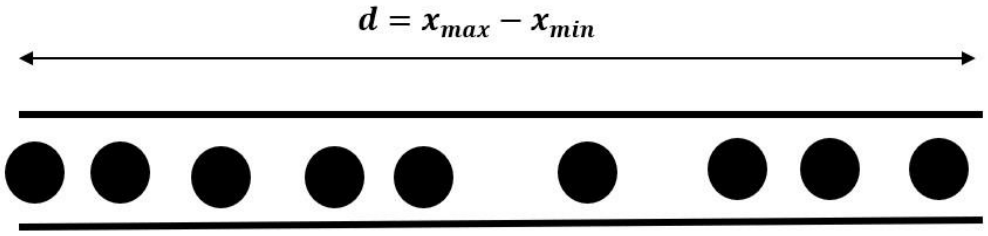

Figure 2. WSN Topology

4.2. Results and Analysis

Extensive simulations are conducted to show the efficiency of the proposed approach, where 1000 topology are randomly generated following the Poisson distribution. Moreover, this is done for different values of PLE and σ, the log-Normal shadowing standard deviation. Simulations are done using MATLAB. All relevant simulation parameters are listed in Table 1.

Table 1. WSN Simulation Parameters Setup

Parameter (Unit)	**Value (s)**
γ: PLE	(3;4)
N: Sensor set cardinality	100
λ: WSN density (average distance between 2 adjacent sensors) (node/m)	$1/3$
σ: Log-Normal shadowing standard deviation (dB)	(1;2;3;4;5;6)
$Pr\ (d_0)$: Received power at reference $d_0 = 1$ (dBm)	-45
P_{th}: Threshold power (dBm)	-90
Anchor number	(32;8;2)
Number of topologies	1000

Positions of anchors are chosen to cover the deployed network, with a step η such as in equation 10.

$$\eta = \frac{x_{max} - x_{min}}{N_A} \tag{10}$$

Where N_A is the number of anchors, unknown position is estimated using multilateration or bilateration in a one -dimensional deployment.

Figures 3 and 4 present the cumulative density function (CDF) of normalized localization error for $\gamma = 3$ for both unknown homogeneous and known homogenous environment respectively for different values of anchor number. With the proposed strategy, until 90% of the sensors could estimate their position with a NE less than 0.04 while using 32 anchors which represent 32% of total node number in an unknown homogeneous WSN. In contrast, 78% of sensors achieve the same accuracy with only 2 anchors, when the WSN is unknown homogeneous a priori. However, 90% of sensors estimate positions with NE equals to 0.1 with 2 anchors.

On the other hand, 90% of the sensors estimate their position with a NE less than 0.01 while using 32 anchors, while 82% of sensors achieve the same accuracy with only 2 anchors in a known homogenous WSN. On the other hand, 90% of sensors estimate the position with an error equal to 0.018 with 2 anchors.

Likewise, figures 5 and 6 present CDF of normalized localization error for$\gamma = 4$. Results in figure 5 where the WSN is unknown homogeneous show that until 90% of the sensors could estimate their position with a NE less than 0.012 using 32 anchors, this percentage decreases to 83% while using only 2 anchors. Moreover, 90% of the sensors could estimate positions with a NE less than 0.02 with 2 anchors. Also, figure 6 shows the same results as 90% of sensors achieve an error of 0.03 with 32 anchors, and 76% of sensors achieve this error with 2 anchors in a known homogenous WSN.

Results obtained show efficiency of the technique in using less anchors while maintaining high localization accuracy. In the example used in this case the anchor's number is passing from using 32 anchors to only 2 anchors with a little increase in error values, $\Delta NE = 0.006$ for $\gamma = 3$ in a priori unknown homogeneous WSN, $\Delta NE = 0.008$ for $\gamma = 3$ in a known homogeneous WSN, $\Delta NE = 0.008$ for $\gamma = 4$ in a priori known homogeneous WSN and $\Delta NE = 0.018$ for $\gamma = 4$ in an unknown homogeneous WSN. Decreasing anchor nodes will decrease network cost which is an important constraint in WSN. In addition to, it can be observed that in homogeneous network the

localization errors are less than those obtained in an unknown homogeneous network. This shows the advantage of knowing a priori that a WSN is homogeneous., i.e., nodes have a priori the same communication range R_i.

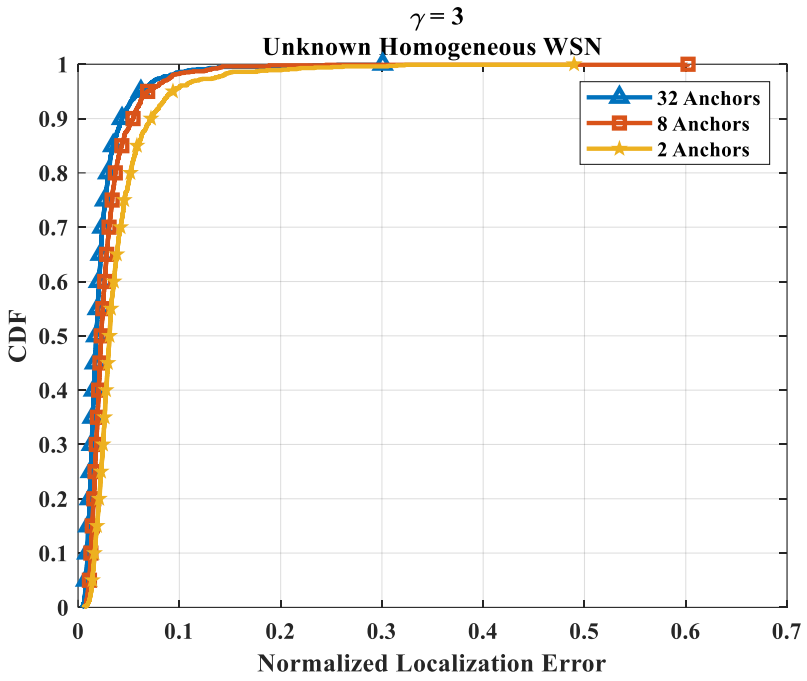

Figure 3. CDF of Normalized Error for PLE=3 in an Unknown Homogenous WSN

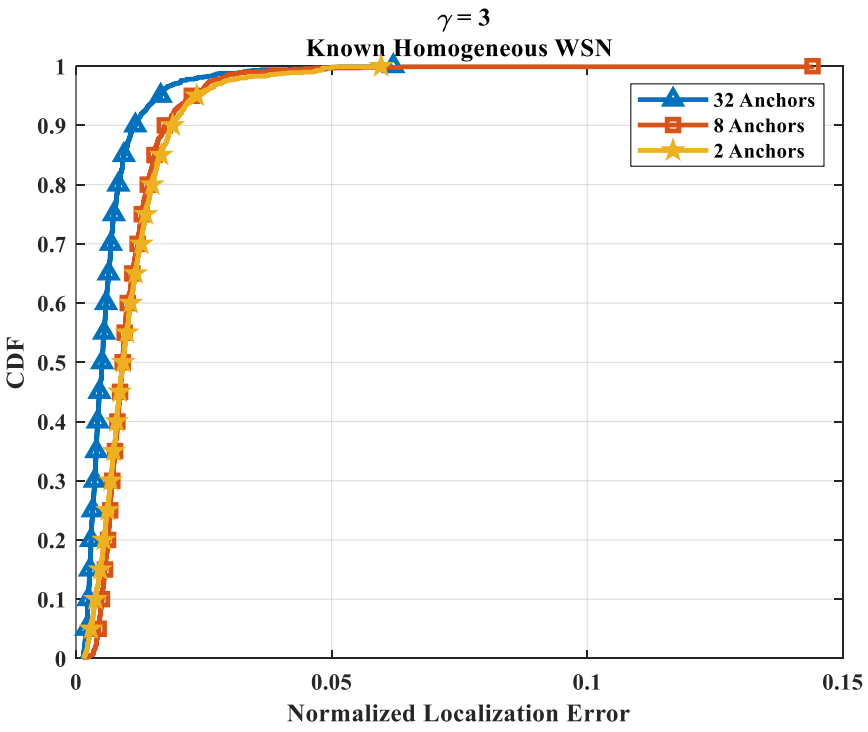

Figure 4. CDF of Normalized Error for PLE=3 in a Known Homogenous WSN

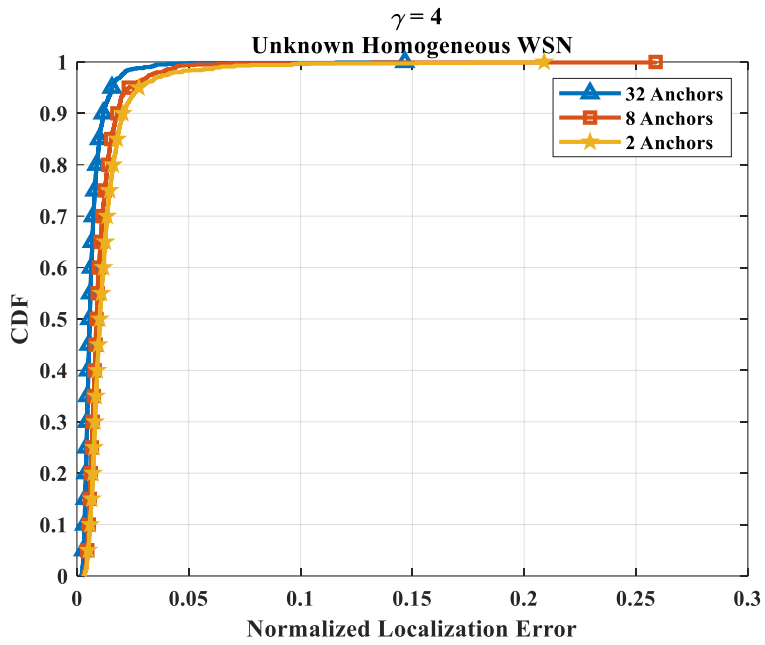

Figure 5. CDF of Normalized Error for PLE=4 in an Unknown Homogenous WSN

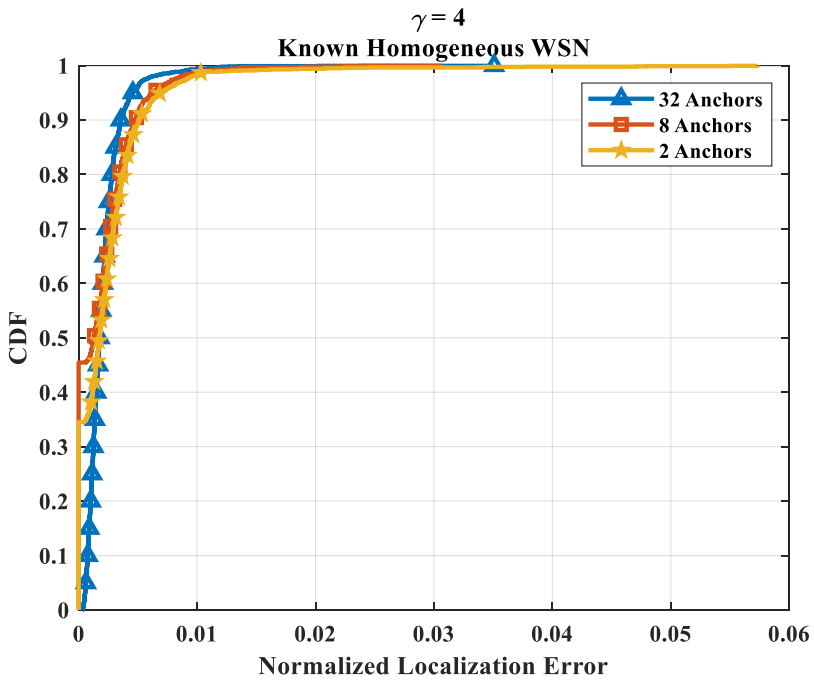

Figure 6. CDF of Normalized Error for PLE=4 in a Known Homogenous WSN

5. CONCLUSIONS

In this paper an anchor number optimization in localization in WSN is presented. By using the approach based on connectivity and network information, the method is able to localize sensors in a WSN with a very low number of anchors and with high accuracy. Hence, the efficiency of the proposed approach based on estimating channel properties to compensate the anchor number in

the localization process is proved. Indeed, in terms of anchor rates results show reduction in anchor deployment rates from 32% to 2%. This solution was derived for one-dimensional WSNs used in many new applications. However, extensions to two- or three-dimensional network topologies are under investigation. Also, other deployment assumptions are under investigation such as gaussian deployment.

REFERENCES

[1] Wang, Zengfeng; Zhang, Hao; Lu, Tingting& Gulliver, T. Aaron, (2018) "A Grid-Based Localization Algorithm for Wireless Sensor Networks Using Connectivity and RSS Rank,"*in IEEE Access*, vol. 6, pp. 8426-8439.

[2] Muhammad Aamir, Cheema, (2018) "Indoor location-based services: challenges and opportunities", *SIGSPATIAL Special* ,Volume 10, Issue 2, pp 10–17.

[3] Ian F. Akyildiz & Mehmet Can Vuran, (2010) *Wireless Sensor Networks*. John Wiley Sons Ltd.

[4] Vijayalakshmi S. R & Muruganand, (2018) *Wireless Sensor Networks: Architecture, Applications, Advancements. Mercury Learning and Information*: Dulles, Virginia.

[5] Leila, Chelouah; Fouzi, Semchedine & Louiza Bouallouche-Medjkoune, (2018) "Localization protocols for mobile wireless sensor networks: A survey", *Computers & Electrical Engineering*, Volume 71, pages 733-751.

[6] Neha Sharma & Vishal Gupta, (2020) "Meta-Heuristic-Based Optimization of WSNs Localisation Problem- a Survey", *Procedia Computer Science*, Volume 173, Pages 36-45.

[7] Tomic, Slavisa; Beko, Marko & Dinis, Rui, (2015) "RSS-Based Localization in Wireless Sensor Networks Using Convex Relaxation: Noncooperative and Cooperative Schemes," IEEE *Transactions on Vehicular Technology*, vol. 64, no. 5, pp. 2037–2050.

[8] Xu, Enyang; Ding, Zhi & Dasgupta, Soura, (2011) "Source Localization in Wireless Sensor Networks from Signal Time-of-Arrival Measurements," IEEE *Transactions on Signal Processing*, vol. 59, no. 6, pp. 2887–2897.

[9] Zhao, Yue; Li, Zan; Hao, Benjian & Shi, Jia, (2019) "Sensor Selection for TDOA-Based Localization in Wireless Sensor Networks with Non-Line-of-Sight Condition," in IEEE *Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9935-9950.

[10] Shao, Hua-Jie; Zhang, Xiao-Ping & Wang, Zhi, (2014) "Efficient Closed-Form Algorithms for AOA Based Self-Localization of Sensor Nodes Using Auxiliary Variables," in IEEE *Transactions on Signal Processing*, vol. 62, no. 10, pp. 2580-2594.

[11] Ru, Lulu & Zhang, Linghua, (2017) "A weighted Centroid Localization Algorithm for Wireless Sensor Networks Based on Weight Correction," 2017 9th *International Conference on Advanced Infocomm Technology* (ICAIT), pp. 165-169.

[12] Wang & Jin H. (2009) "Improvement on APIT localization algorithms for wireless sensor networks," in *Proceedings of the International Conference on Networks Security, Wireless Communications and Trusted Computing* (NSWCTC '09), pp. 719–723.

[13] Li, Xiaoyin; Yan, Lianshan; Pan, Wei & Luo, Bin (2017) "Secure and Robust DV-Hop Localization Based on the Vector Refinement Feedback Method for Wireless Sensor Networks," in *The Computer Journal*, vol. 60, no. 6, pp. 810-821.

[14] Ibrahim, Nemer; Tarek, Sheltami; Elhadi, Shakshuki; Abdulrahman, Abu Elkhail&Mumin Adam, (2021) Performance evaluation of range-free localization algorithms for wireless sensor networks. *Pers UbiquitComput* 25, pp. 177–203.

[15] Nour, Zaarour; Sofiène, Affes; Nahi, Kandil & Nadir, Hakem, (2020) "Connectivity-Based Joint Parameter Estimation in One-Dimensional Wireless Sensor Networks," 2020 *International Wireless Communications and Mobile Computing* (IWCMC), pp. 358-364.

[16] Dai, Wenhan; Shen, Yuan& Win, Moe Z Dai W, (2012) "On the minimum number of active anchors for optimal localization," 2012 *IEEE Global Communications Conference* (GLOBECOM) pp. 4951-4956.

[17] Du, Shifeng; Huang, Baoqi; Jia, Bing & Li, Wuyungerile, (2019) "Optimal Anchor Placement for Localization in Large-Scale Wireless Sensor Networks," IEEE *5th International Conference on Data Science and Systems* (HPCC/SmartCity/DSS), pp. 2409-2415.

[18] Lasla, Noureddine; Younis, Mohamed; Ouadjaout, Abdelraouf & Badache, Nadjib, (2015) "On optimal anchor placement for efficient area-based localization in wireless networks," 2015 *IEEE International Conference on Communications* (ICC), pp. 3257-3262.
[19] Ammari, Habib M., (2014) *The Art of Wireless Sensor Networks*: Volume 1 Fundamentals. Springer Publishing Company, Incorporated.
[20] Savvides, A; Garber, W.L.; Moses, R.L. & Srivastava, M.B, (2005) "An Analysis of Error Inducing Parameters in Multihop Sensor Node Localization". IEEE *Transactions on Mobile Computing*, 4:567–577.
[21] Shah, Kaushal & Jinwala, Devesh C., (2016) "A Secure Expansive Aggregation in Wireless Sensor Networks for Linear Infrastructure," in Proc. *of IEEE Region 10 Symposium (TENSYMP)*, Bali, Indonesia, pp. 207-212.
[22] Rappaport T.S., (1999)*Wireless Communications*. New Jersey: Prentice Hall.
[23] Isabelle, Vin; Davy, P. Gaillot; Pierre, Laly; Martine, Liénard & Pierre, Degauque, (2015) "Overview of Mobile Localization Techniques and Performances of a Novel Fingerprinting-Based Method", *Comptes Rendus Physique*, Volume 16, Issue 9, Pages 862-873.
[24] Haenggi, Martin; Andrews, Jeffrey G.; Baccelli, Francois; Dousse, Olivier & Franceschetti, Massimo, (2009) "Stochastic geometry and random graphs for the analysis and design of wireless networks," in IEEE *Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029-1046.
[25] Jesper, Moller & Rasmus, Plenge Waagepetersen, (2004) *Statistical Inference and Simulation for Spatial Point* Process, Chapman and Hall/CRC.

AUTHORS

Nour Zaarour: received the Bachelor's Degree in biomedical engineering from the Islamic University of Lebanon (IUL), Khaldeh, Lebanon in 2010, and the M.Sc. degree from the University of Quebec in Abitibi-Témiscamingue (UQAT), Val-d'Or, Qc, Canada, in 2013. She is currently pursuing the PhD. degree with the UQAT. Her research interests include channel modeling, localization algorithms, wireless sensor networks and neural networks.

Nadir Hakem is in charge of innovation and digital strategy for Meglab a subsidiary Epiroc company, Canada since May 2021. Previously, he was Professor at the School of Engineer from 2009 to 2021 and he has been a research assistant from 2006 to 2009 at the University of Québec Abitibi-Témiscamingue (UQAT). He is involved in research work in radio frequency measurements and analytical modeling, energy harvesting and localization issues for wireless sensor network.
Dr. Hakem obtained his Advanced Studies Diploma and his PhD in Computer Science degree respectively in 1999 and 2004 both at Blaise Pascal University, France. He holds an engineer's degree in computer science from the science and Technology University of Algiers, Algeria. He is also a member of the Order of Engineers of the Province of Québec (OIQ) and the Institute of Electrical and Electronics Engineers (IEEE). He has been involved in many scientific conference committees and in supervision of over 30 graduate and post-graduate students.

NahiKandil is a professor at the School of Engineering in the University of Quebec in Abitibi-Timiskaming (UQAT) since 2000. He has worked on research dealing with neural networks and their applications to wireless communications and power systems. Prof. Kandil is now working mainly with localization and channel modeling in confined areas and mine environments.

PREDICTIVE MODELLING OF COVID-19 STIMULUS FUNDS PAID FOR NURSING HOME QUALITY INCENTIVE PROGRAM

Omar Al-Azzam and Paul Court

Department of Computer Science and Information Technology,
St. Cloud State University, St. Cloud, MN 56301, USA

ABSTRACT

Painstaking measures should be taken to determine how federal dollars are spent. Proper justification for allocation of funds rooted in logic and fairness leads to trust and transparency. The COVID-19 pandemic has warranted rapid response by government agencies to provide vital aide to those in need. Decisions made should be evaluated in hindsight to see if they indeed achieve their objectives. In this paper, the data collected in the final four months of 2020 to determine funding for nursing home facilities via the Quality Incentive Program will be analysed using data mining techniques. The objective is to determine the relationships among numeric variables and formulae given. The dataset was assembled by the Health Resources and Services Administration. Results are given for the reader's insight and interpretation. With the data collection and analytical process, new questions come to light. These questions should be pondered for further analysis.

KEYWORDS

Predictive modelling, Cross validation, Linear Regression.

1. INTRODUCTION

Over the span of approximately a year, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a virus that has come to be known as COVID-19, emerged as a major health concern for people over the entire globe [1]. Oddly, older people are disproportionately affected by its adverse consequences. These devastating effects have placed an inordinate financial burden on congregate care facilities for the elderly. Roughly eighty percent of COVID- 19 deaths in the United States have been people ages 65 and older. A person 85 years of age or older has a rate of death 8700 times higher than a person 5 – 17 years of age [2]. For this reason, government funding of nursing home care facilities is of vital importance. The Department of Health and Human Services provided $2 billion as incentive payments to nursing home facilities that found techniques to lower COVID-19 infection rates and facility mortality. The Coronavirus Aid, Relief, & Economic Security (CARES) Act, a bipartisan group, collaborated with the Paycheck Protection Program and Health Care Enhancement Act (PPPHCEA) and the Coronavirus Response and Relief Supplemental Appropriations (CRRSA) Act to form an alliance tasked with determining a fair method of allocating relief funds to hospitals, nursing homes, and other front line health care facilities. Their express purpose was to assist with coronavirus expenses incurred by these facilities. These procedures would be deemed a success if the facility were able to demonstrate reduction relative to their facility's county infection rate and mortality versus national metrics [3][4].

David C. Wyld et al. (Eds): CSTY, SIGI, MaVaS - 2021
pp. 13-25, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.112202

1.1. The Research

This research will discuss an analysis of the Nursing Home Quality Incentive Program (QIP). The program was designed for the purpose of defending nursing home patients across the country against severe outcomes due to the ongoing pandemic. The program designed a method of allocating funds to facilities based on their performance [3][5]. Mortality rates that significantly exceeded the national average in each month could not receive payment for that month. To ensure fairness in the process of allocating funds via QIP, the facility needed to meet the eligibility requirements for the performance period. A metric was created that involves infection rate and mortality rate to those facilities that followed the guidelines for QIP data reporting [5]. The specific data gathered to determine QIP funding will be discussed in detail and relationships among data will be presented to detect any anomalies in funding. The evidence presented will help to make future decisions about funding patterns and processes more astutely.

1.2. The analysis processes

When data scientists are faced with predicting results for a given dependent variable, the usual process is to examine each data variable individually for shape, centre, and variability, noting any curious results that may expose themselves. Next, pairing data variables is done to evaluate strength of fit, association, and the nature of the relationship. Strength and association can be measured using correlation (positive and negative association), but the nature of the relationship (linear or non-linear) can be difficult to determine superficially [6]. Multivariate evaluations combine the effects of independent variables on a dependent variable, usually for the purpose of predicting future outcomes. Multiple independent variable predictions are often made either by ordinary least squares or cross validation techniques [7]. Figure 1 depicts this type of process. Once an iteration of the process has been completed, further investigation is usually warranted to study associations or anomalies detected in the previous cycle.

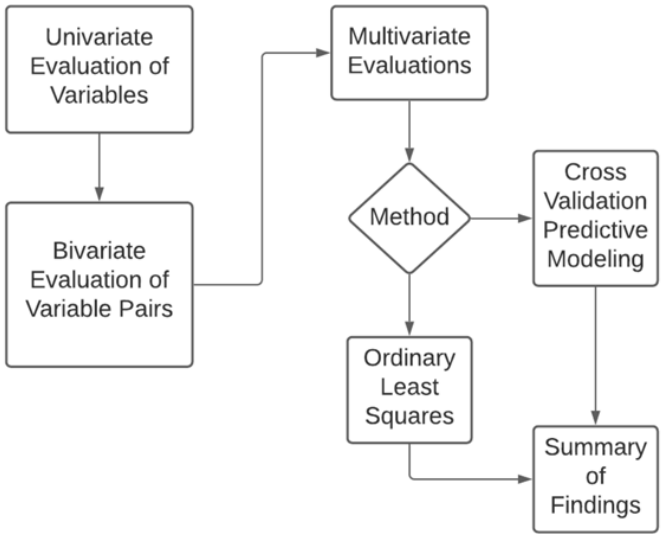

Figure 1. A standard evaluation process for critiquing datasets for the purpose of predicting values for a dependent variable. This investigative process can (and is often) repeated as anomalies are discovered.

1.3. Dataset information

The data studied in this research has 33,305 entries from across the country. Among the fifteen attributes noted are CCN number, Facility Name, City, State, Zip Code, Total Resident Weeks (TRW), Total COVID Infections (TCI), Facility Infection Rate Per 1000 Resident Weeks (FIR), County Infection Rate Per 1000 Resident Weeks (CIR), Infection Performance Score (IPS), Infection Performance Score Capped (IPSC), Mortality Adjustment (MA), Performance Month (PM), and Final Payment (FP) [3].

The CCN number is known as the Centers for Medicare and Medicaid Services certification number, sometimes called the Medicare Provider Number. It is a facility identification number used by government agencies. After the identification of the facility and its location, the numeric data in the report are a described in further detail. The data value Total Resident Weeks (TRW) is the number of residents reported in a performance period using the sum of total beds occupied reduced by the number of COVID admissions in each week of the performance period. A facility that meets the requirements will be assigned a performance score for that performance period (one month).

1.4. Calculations for Infection Performance Score and Mortality Adjustment

The steps that follow briefly summarize the performance score calculation:

Step 1: The in-facility infections in a performance period are determined by summing each week's infections as reported.
Step 2: The number of resident weeks is the sum of the total beds occupied in a facility reduced by the number of reported COVID admissions in each week of the performance period.
Step 3: The facility infection rate for a performance period is the ratio of the infections reported in a facility to the total resident weeks as calculated in steps 1 and 2.
Step 4: The county infection rate is a sum of the ratio of infections reported each week, to the total county resident weeks.
Step 5: Assuming the facility infection rate mirrors the county infection rate, an expected number of facility infections is calculated.
Step 6: Finally, the difference between the estimated and the actual infections is calculated. The results are the infection performance score for the given facility in the given county [5]. Formulae for these calculations can be found in the referenced materials.

The infection performance score provides a metric to determining how well a facility is performing versus their county and other facilities throughout the nation.

In much the same way, the mortality adjustment calculates how well a facility performs. Using county mortality estimates, an expected facility mortality is calculated. A score is then assigned to represent the difference between expected and actual deaths.

2. LITERATURE REVIEW

As mentioned, COVID-19 has disproportionally affected the lives of older Americans. Less than one percent of the Americans live in what could be categorized as a long-term care facility. However, residents and facility employees make up about forty percent of COVID-19 deaths [8]. Research about this topic indicates mixed opinions on the effectiveness of formulae applied to the data collected for the purpose of assigning funding to long-term care facilities. Usually, flat fees are charged, but recently, extra fees for specialty services have changed revenue patterns. "These

performance payments are an important stimulus for nursing homes fighting to improve their performance in a dire situation," Terry Fulmer, president of the John A. Hartford Foundation and a member of a commission on coronavirus safety in nursing homes. "As we approach the rollout of safe and effective vaccines for our most vulnerable, we continue the innovative program we created this year to incentivize and assist nursing homes in battling COVID-19 and applying the right infection control practices," said HHS Secretary Alex Azar. "This half a billion dollars in incentive payments will reward nursing homes that have shown results in their tireless work to keep their residents safe from the virus." [9]

States that have controlled the virus well in their communities, however, would typically be placed at a funding allocation disadvantage. As a congressional delegation noted, "The fact that there is a lower level of COVID-19 spread in the community in New Hampshire does not mean that Granite State nursing facilities do not need support. That is why incorporating measures of overall community spread of COVID-19 (outside of nursing facilities) into the formula is so damaging for states like New Hampshire." [10] In another example, Wisconsin's nursing home population comprises less than two percent of the nation's nursing home population but are receiving more than four percent of the emergency funds. October saw the state's positivity rate soar, making it easier for facilities to perform well with the QIP metrics [9].

Mathematicians have attempted to model natural occurring events, such as the spread of disease. This pandemic is no exception. Models can help predict the impact of these events not only by estimating cases and mortality, but necessities such as peak need for hospital beds. For a prediction to be effective, it must consider infection rate of detected and undetected cases, number of susceptible people in a population, along with those who are immune to infection. These models need to consider parameters that cannot be practically measured [11]. Predictive models should consider how public assistance can be most effectively allocated. Prognostications should tie economics to the instance of disease and the extent to which the virus has impacted individual's financial health. Factors should be considered to manage such predictions on economic impact include lockdown measures, job loss, health related expenses, socio-economic status, ethnicity, loan eligibility, even social distancing [12].

3. METHODOLOGY

To better understand how Nursing Home Quality Incentive Program funds have been distributed, its dataset will be analysed diligently. A summary of the numeric data collected for purpose of distributing funds will be organized and evaluated. This will be followed by paired comparisons between variables to see if associations exist. The relationship between pairs of variables will be evaluated using correlation as the metric. These relationships will help to narrow the focus of an ordinary least squares regression analysis with multiple independent variables. The regression equation will help interpret whether funding allocations are justified. Finally, a ten-fold cross validation procedure will be done to test the model created for the final payment of funds.

3.1. Conjectures

A conjecture would be that facility infection rate per 1000 resident weeks would be positively related to county infection rate per 1000 resident weeks. Also expected is a positive relationship between facility infection rate per 1000 resident weeks and the infection performance score. Lastly, a positive association with facility and county infection rate per 1000 resident weeks and final payment or infection performance score and final payment should be found. This would indicate a justified method for assigning funding to the facilities. High positive correlations among any of these variables of interest would not be surprising. One of the most common

methods for calculating correlation is the Pearson's correlation coefficient. It is found by summing the distance each data item is from the mean then dividing by the standard deviation [6]. For instance, there would be a high degree of positive correlation between overtime worked and money earned by an employee. If hours of overtime worked increases, there should be a somewhat predictable increase in money earned.

Equation 1. Pearson's Correlatioin Coefficient.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

The variables, however have to have a justifiable linear association for correlation to be relevant. Correlation, by nature of its calculation will always be between negative one and one, with negative one or one being perfect correlation. A high positive correlation for this scenario would be $r = 0.85$ or higher, indicating the variables are strongly associated. The associations will be analysed further by taking the data in monthly segments. If monthly correlations increase or decrease, the rationale for these changes will be considered.

3.2. Outline of the process

With these preliminary activities complete, ordinary least squares regression will consider the combined effects of the six major independent variables in the dataset: Total Resident Weeks, Total Covid Infections, Facility Infection Rate, County Infection Rate, Infection Performance Score, and Mortality Adjustment. These multiple independent variables will be used to create a linear regression model to predict Final Payment, the dependent variable. The goodness of fit for this model will be evaluated using correlation coefficient and probability models.

Next, predictive modelling using cross validation methods will be examined. *K*-fold cross validation consists of partitioning data into disjoint groups. One data segment is held out as the testing partition, developing the prediction model using the remaining partitions as a training set. When the model is trained, it can be compared to the actual data in the unused partition. Repeated *K* times and combined, an overall model will be produced and evaluated.

Other tools used to evaluate the relationship among the variables will include graphical analysis. We must look at the data to help understand them. This will assist in detecting underlying dimensions in the data.

4. RESULTS

A summary of the numeric data collected in the last four months of 2020 by the Centers for Disease Control and Prevention will be followed by pairing variables with Final Payment. Linear regression analysis with correlation as the metric will be used to evaluate the relationships.

4.1. Univariate evaluation

Figure 2 summarizes the state-by-state frequencies of nursing home facilities involved in the Provider Relief Fund COVID-19 Nursing Home Quality Incentive Program. All fifty states are represented as well as the District of Columbia and Puerto Rico.

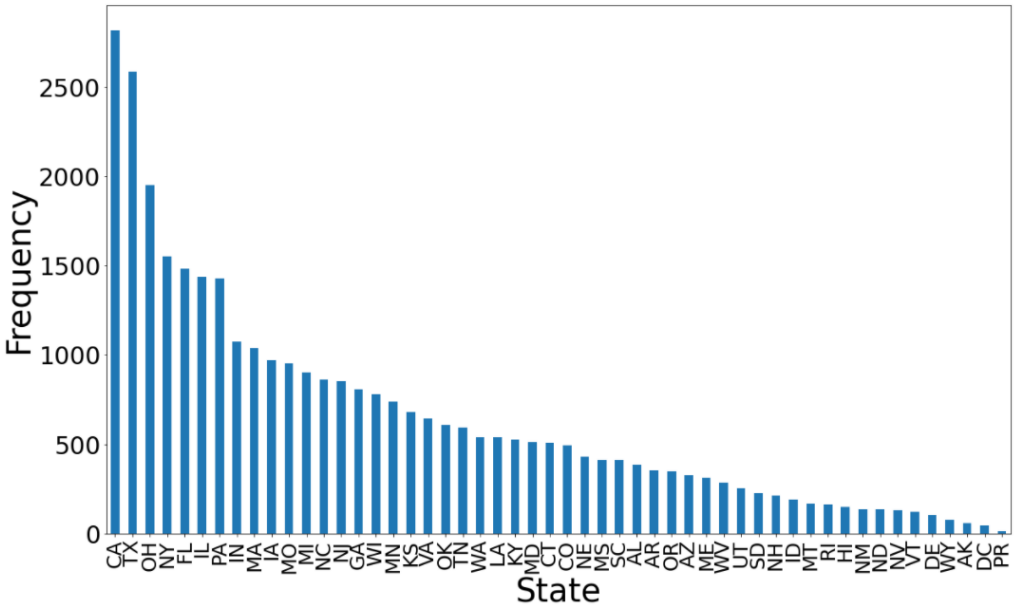

Figure 2. State by state, including District of Columbia and Puerto Rico, enrolment of nursing home facilities in the QIP. Each state with DC and PR listed on the horizontal axis and the number of facilities in each state on the vertical axis.

A histogram of the primary numeric variables represented in the dataset can be found in Figure 3. Each distribution shows a distinct skew. The top left, Total Resident Weeks is the least skewed, whereas Facility Infection Rate per 1000 Resident Weeks is the most skewed. The similar shapes would lead to the belief that these data are highly correlated.

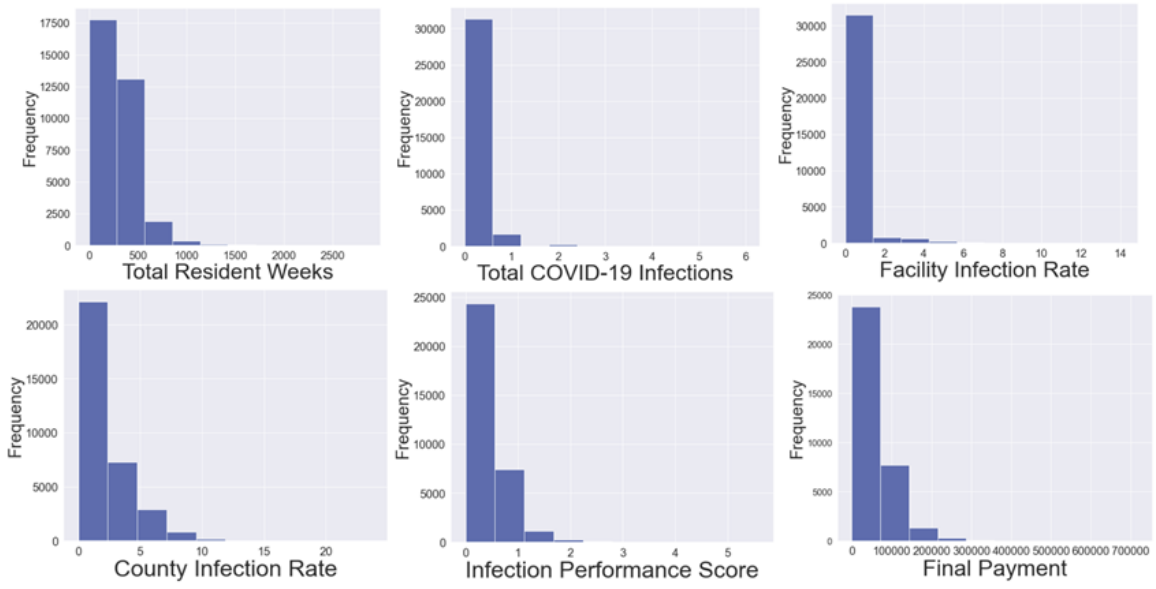

Figure 3: Top left; histogram of Total Resident Weeks (*x*-axis show TRW raw numbers vs. frequency of occurence in the data set on the *y*-axis), Top middle; histogram of Total Covid Infections (*x*-axis show TCI raw numbers vs. frequency of occurence in the data set on the *y*-axis), Top right; histogram of Facility Infection Rate per 1000 Resident Weeks (*x*-axis show FIR in rate per 1000 resident weeks vs. frequency of occurence in the data set on the y-axis), Bottom left; County Infection Rate per 1000 Weeks (*x*-axis show CIR rate per 1000 resident weeks vs. frequency of occurence in the data set on the *y*-axis), Bottom middle;

Infection Performance Score (*x*-axis show IPS vs. frequency of occurence in the data set on the *y*-axis), Bottom right; Final Payment (*x*-axis in Dollars vs. frequency of occurrence in the data set on the *y*-axis).

A numeric snapshot of each of the variables is provided in Table 1. It depicts in the dataset for the Provider Relief Fund COVID-19 Nursing Home Quality Incentive Program. The skewing of the data is again evident. The mean of each data set is larger than the median, except Infection Performance Score Capped. The minimum and maximum values for each variable also add depth to the variability and amount of skew present.

Table 1. A snapshot of statistics for the variables in the dataset. Note the mean and median of each. All are skewed to the right except Infection Performance Score Capped.

	TRW	**TCI**	**FIR**	**CIR**	**IPS**	**IPSC**	**MA**	**FP**
Mean	307.73	0.07	0.20	2.20	0.43	0.73	0.06	57,812.48
Std	195.50	0.31	0.87	2.02	0.37	0.44	0.09	49,371.35
Min	1.00	0.00	0.00	0.01	0.00	0.00	-0.2	100.67
Med	273.00	0.00	0.00	1.48	0.34	1.00	0.00	45,629.29
Max	2849.00	6.00	14.18	23.78	5.60	1.00	0.20	718,593.32

4.2. Bivariate evaluation

A heatmap of correlations among the numeric variables in the QIP data is shown in Figure 4. It compares the variables two at a time and displays the Pearson's correlaion coeffeiceient for each pair. The heatmap is symmetric with the diagonal comparing the data variable to itself and thus creating a correlation of perfect fit, $r = 1.0$. It can be disregarded. The calculations shown in Figure 4 were done using Jupyter/scipy stats package to calculate the Pearson's correlation coefficient in the standard way. The formula assumes a linear realtionship between the bivariate data and represents the ratio of the covariance of the variables to the product of their standard deviations. The heatmap shows surprising results. Among them, Final Payment funding is almost perfectly correlated with the Infection Performance. Total Covid Infection and Facility Infection Rate per 1000 Resident Weeks ranked a distant second among the relationships analysed. Final Payment vs. Total Resident Weeks, Total Resident Weeks vs. Infection Performance Score and County Infection Rate per 1000 Resident Weeks vs. Infection Performance Score round out the top five highly correlated bivariate relationships, the last having a correlation coefficient of 0.42. Referring to the original assumptions, facility infection rate per 1000 resident weeks would be positively related to county infection rate per 1000 resident weeks (true, however $r = 0.39$). Also expected is a positive relationship between facility infection rate per 1000 resident weeks and the infection performance score (true, but $r = 0.038$). Lastly, a positive association with facility infection rate and county infection rate per 1000 resident weeks and final payment ($r = 0.058$ and $r = 0.42$, respectively) should be found. This would indicate a justified method for assigning funding to the facilities.

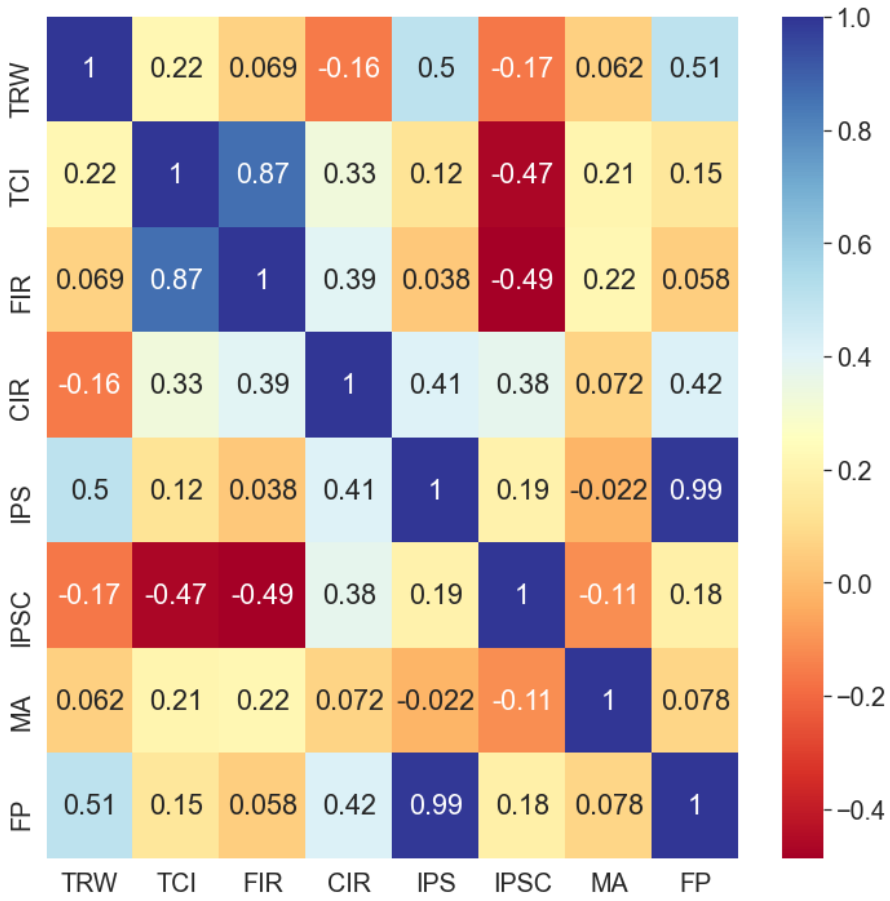

Figure 4: Heat map of numeric data bivariate correlations. The diagonal shows the variable's relationship with itself (r = 1.0). Most correlations are not very strong ($|r| < 0.50$). However, Final Payment with Infection Performance Score (r = 0.99) and Facility Infection Rate with Total Covid Infection (r = 0.87) show that these variables could have linear association.

When each of the top five relationships are investigated more closely, the month-by-month relationships showed interesting results as captured in Table 2 and depicted graphically in Figure 5. If compared to the overall correlation, the monthly correlation for final payment and infection performance score along with total covid infection and facility infection rate per 1000 resident weeks remained consistent. However, there were noticeable increases in correlation between final payment and total resident weeks, indicating that facilities with larger populations received more funding. Total resident weeks and infection performance score also showed increased correlation over the four-month period.

Table 2. An analysis of related variables separated by month

Relationship	Overall	Sept	Oct	Nov	Dec
FP vs IPS	0.99	0.99	0.99	0.99	0.99
TCI vs FIR	0.87	0.94	0.88	0.85	0.87
FP vs TRW	0.51	0.36	0.39	0.75	0.83
TRW vs IPS	0.50	0.35	0.38	0.75	0.83
CIR vs IPS	0.41	0.74	0.65	0.07	0.02

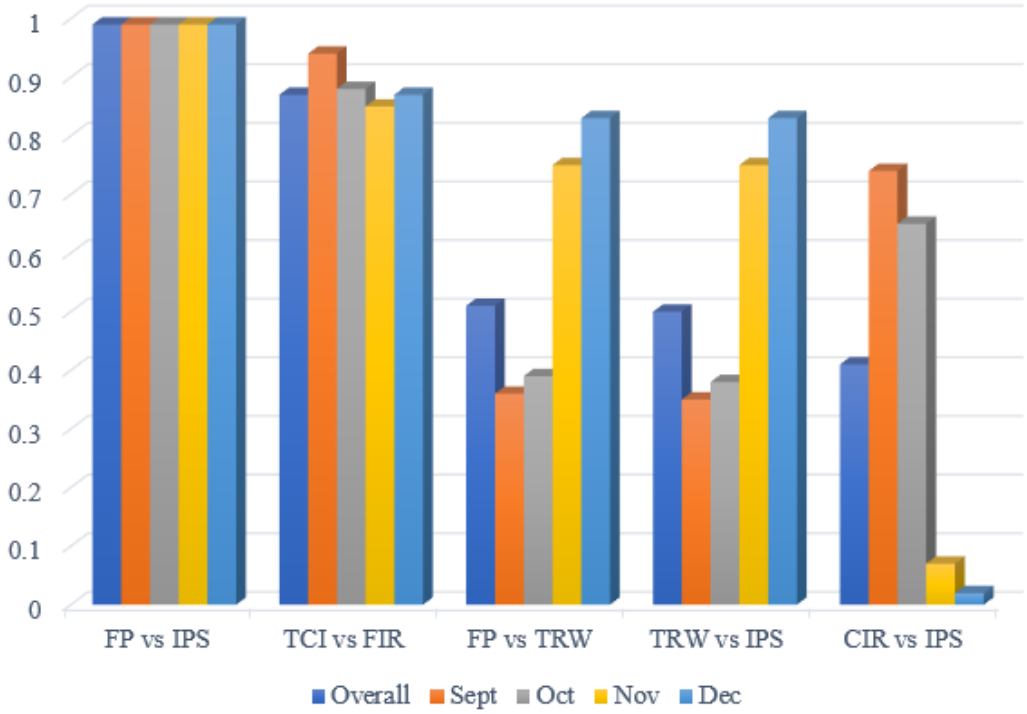

Figure 5. Graphical representation of the analysis of related variables separated by month. Overall correlation is followed by monthly results.

County infection rate per 1000 resident weeks and infection performance score correlation weakened significantly, which was a goal of the program. Recall that performance scores for a successful facility is based on having facility infection rates that do not reflect the county facility rate. However, note a decline in facility participation in the QIP over this period due to failure to meet program requirements tied to under performance.

4.3. Check for normality

Since each of the data was skewed when analysed graphically and statistically, a logarithmic transformation was applied to examine whether the data could be normalized. The results are summarized in Figure 6. If the dataset could have zero as a value, it was adjusted so the transformation could be performed. The transformation normalizes Total Resident weeks and Final Payment quite well and County Infection rate and Infection Performance score reasonably well. Total Covid Infections and Facility Infection Rate were not normalized by the logarithmic transformation. If the assumption made previously is correct, there should be a high degree of association between TRW and FP, some association between CIR and IPS, and the relationship between TCI and FIR should be weak. A formalized hypothesis regarding the relationships found in this data set will reflect the assumptions made earlier.

4.4. Formal hypotheses

Each of these hypotheses will be tested against the following two-tailed null hypothesis:

H0: The relationship expected is not present in the data.
H1: County infection rate will have a positive impact on facility infection rate.
H2: Facility infection rate will have a positive influence on infection performance score.

H3: Facility infection rate and county infection rate will dictate a facility's infection performance score and therefore their final payment.

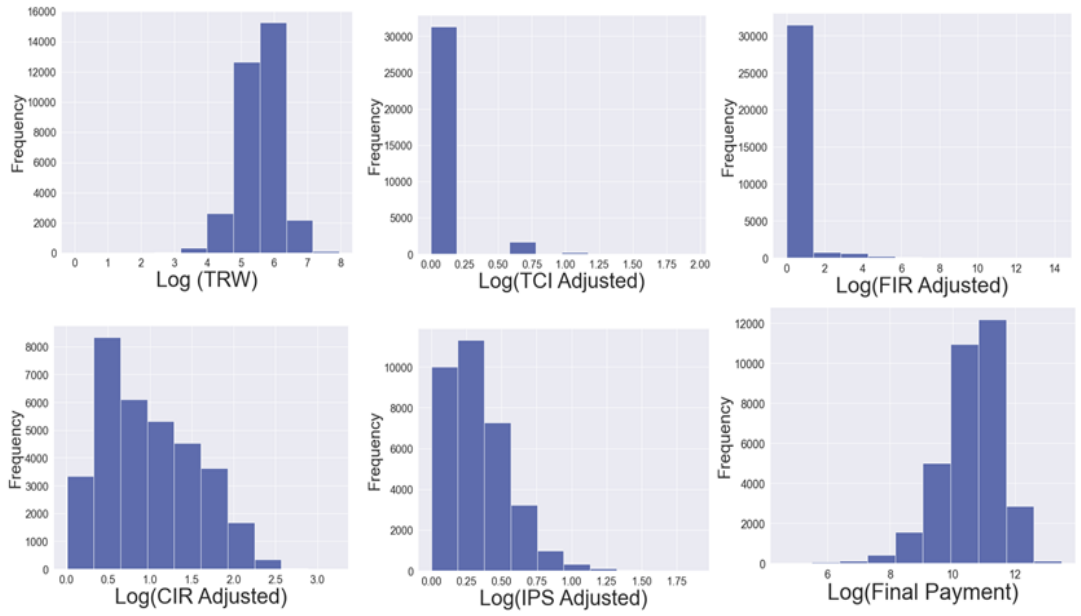

Figure 6: Top left; frequencies vs. log(Total Resident Weeks), Top middle; frequencies vs. log(Total Covid Infections), Top right; frequencies vs. log(Facility Infection Rate per 1000 Resident Weeks), Bottom left; frequencies vs. log(County Infection Rate per 1000 Weeks), Bottom middle; frequencies vs. log(Infection Performance Score), Bottom right; frequencies vs. log(Final Payment). All figures show a histogram with the frequencies on the *y*-axis and logarithmic transformations to numeric data variables found in Figure 2 on the *x*-axis. Data sets with a zero element were adjusted to allow a logarithmic transformation.

4.5. Multivariate evaluation

Using ordinary least squares regression with multiple independent variables Total Resident Weeks, Total Covid Infections, Facility Infection Rate, County Infection Rate, Infection Performance Score, and Mortality Adjustment a model to predict the Final Payment was produced. The resulting coefficients and measures of evaluation are found in Table 3. The *t*-values help to show the statistical significance of each coefficient. Probabilities under the heading $P > |t|$ give the likelihood of a result of this nature occurring by chance, assuming the two-tailed null hypothesis to be true. If the *P*-value is less than the confidence level (usually $\alpha = 0.05$), it indicates a statistically significant result [13]. Facility Infection Rate has a negative impact on the Final Payment. The *t*-value of -18 suggests that the assumption that FIR has no influence on FP is extremely unlikely. The coefficient for FIR is -990.130, indicating that its values detract from the Final Payment prediction. County infection rate also has a negative impact (t = -3.153 with coefficient -49.026), but in a much smaller magnitude. Infection Performance Score had an extreme positive impact on the prediction of the Final Payment (t = 1495.260 with coefficient 134,500). Total Covid Infection also had a positive effect on Final Payment (t = 19.867 with coefficient 2945.970), but in a much smaller way, in comparison. The coefficient of determination (r^2 = 0.993) for the model shows the proportion of the variance in Final Payment that can be attributed to the multiple independent variables assessed. Standard errors for each variable were calculated using the basic standard error of the estimate of the coefficient as provided by the ordinary least squares regression feature in python.

A ten-fold cross validation procedure was applied to the data, partitioning the data into ten disjoint segments. The dependent variable choice for the models was again Final Payment. The independent variables used were the same as in the ordinary least squares calculation. Models were trained on nine of the folds, using the tenth to evaluate the model's skill.

Table 3. Coefficients of the ordinary least squares regression model with probability values for each independent variable

	Coefficient	**Standard Error**	***t*-value**	***P* > |*t*|**
Intercept	-2465.098	55.510	-44.408	0.000
TRW	-0.981	0.160	-6.152	0.000
TCI	2945.970	148.282	19.867	0.000
FIR	-990.130	54.284	-18.240	0.000
CIR	-49.026	15.550	-3.153	0.002
IPS	1.345e+05	89.920	1495.260	0.000
MA	5.403e+04	246.312	219.349	0.000

This process was repeated using each of the ten disjoint folds as the test partition. The combined validation score for the ten folds was 0.9891. This number estimates the skill of the model in its overall performance on the test folds. It is expected that the fitted model will perform better on some folds than others. The overall results in this case are very consistent. Table 4 displays the cross-validation score for each of the ten train-test splits.

Whenever regression equations are used, a necessary procedure is to check for interesting residual results. In this case, a plot of the predicted final payment versus the residual values showed a cone shape, indicating that as payments increased, variability increased. However, when examined on a per dollar basis, the smaller final payments showed more variability.

Overall, the ordinary least squares regression and cross validation evaluations show that the infection performance score justifies the final payment allocation statistically. The methodology for determining the infection performance score may be a topic for debate for future decision-makers, but the final payments received by facilities statistically adheres to the infection performance scores calculated.

Table 4. Cross validation scores for the ten-fold cross validation procedure performed. The *k*-value indicates which of the ten disjoint folds was used to test the data, using the remaining nine as the training data

k	**Validation Score**
1	0.9880
2	0.9857
3	0.9898
4	0.9934
5	0.9938
6	0.9924
7	0.9905
8	0.9934
9	0.9800
10	0.9839

When Final Payment is predicted using the variables not contrived via the formulas provided (IPS and MA), there is a moderate positive correlation (r = 0.546). The positive contributors to Final Payment are Total Covid Infections and Total Resident Weeks with Facility Infection Rate

as the largest detractor. County infection rate was positive, but its coefficient was essentially zero. These results tend to support the design of the infection performance score and the mortality adjustment.

5. CONCLUSIONS

This analysis of the Quality Incentive Program's funding allocation to nursing home facilities across the nation examined the numeric variables and formulae involved. As was demanded by the severity of the pandemic, a rapid response was necessary. Proper evaluation of the response measures is always a good practice to help determine whether a program's goals were met and to help with future decisions about funding of this nature. The program goals were to derive a system to distribute funding tied to facility's performance versus infections and mortality in their surrounding communities. As seen, the infection performance score and final payment are very strongly correlated, providing a transparent method of allocating funds. The data studied for the four-month period also showed successful facilities did manage to out-perform their communities in disease infection and mortality. To improve the quality of the discussion started in this research, data over a longer period would help to determine if trends in funding distributions could be noted. Also of interest, would be performance of facilities that were disqualified from participation in the program as compared to their counterparts who were able to remain in the program. This comparison would shed light on whether the program indeed met the goal of assisting facilities through the pandemic. An extensive inspection from numerous viewpoints is necessary when dealing with such a complex scenario. The results are given for the reader's insight and interpretation. As usual in data analytics, more questions are unearthed that may pique interest and merit further study.

ACKNOWLEDGEMENTS

The authors would like to thank SOLAP Interactive Visualization Platform (grant 211246) for making this research possible.

REFERENCES

[1] World Health Organization, "Naming the coronavirus disease (COVID-19) and the virus that causes it", *WHO*. Accessed on: March 14, 2021. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it

[2] Center for Disease Control and Prevention, (Apr. 16, 2021) "Older adults at greater risk of requiring hospitalization or dying if diagnosed with COVID-19", *CDC*. Accessed on: Apr. 26, 2021. [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/older-adults.html

[3] Health Resources and Services Administration, (2020) "Provider Relief Fund COVID-19 Nursing Home Quality Incentive Program." *DATA.gov*. Accessed on: Mar. 14, 2021. [Online]. Available: https://catalog.data.gov/dataset?publisher=HRSA

[4] Department of Health and Human Services, (Oct. 28, 2020) "Trump Administration distributes incentive payments to nursing homes curbing COVID-19 deaths and infections", *HHS*. Accessed on: February 10, 2021. [Press Release]. [Online]. Available: https://www.hhs.gov/about/news/2020/10/28/trump-administration-distributes-incentive-payments-to-nursing-homes-curbing-covid-19-deaths-and-infections.html

[5] Department of Health and Human Services, (Dec. 7, 2020) "Nursing home quality incentive program methodology", *HHS*. Accessed on: Mar. 18, 2021. [Online]. Available: https://www.hhs.gov/sites/default/files/nursing-home-qip-methodology.pdf

[6] Stojiljkovic, Mirko, "NumPy, SciPy, and Pandas: Correlation with Python", *Real Python*. Accessed on: Mar. 27, 2021. [Online]. Available: https://realpython.com/numpy-scipy-pandas-correlation-python/#example-numpy-correlation-calculation

[7] Zakeri, Z., Mansfield, N., Sunderland, C., and Omurtag, A., (2020) "Cross-validation models of continuous data from simulation and experiment by using liner regression and artificial neural networks." *Elsevier*. Volume 21.

[8] Udow-Phillips, M. and Rontal, R. (2020) "Reforms needed after systemic flaws in nursing homes worsen outcomes from COVID-19." *American Bar Association.* Volume 41, Issue number 6. Accessed on: Feb. 4, 2021. [Online]. Available: https://www.americanbar.org/groups/law_aging/publications/bifocal/vol-41/vol-41--issue-no-6--july-august-2020-/systemic-flaws--tragic-outcomes/

[9] Englund, W., (Dec. 9, 2020) "For the first time, the U.S. will reward nursing homes for controlling the spread of infectious disease", *Washington Post.* Accessed on: Mar. 8, 2021. [Online]. Available: https://www.washingtonpost.com/business/2020/12/09/nursing-home-infection-control/

[10] Williams, B., (Dec. 8, 2020) "Nursing home quality incentive program? Hardly", *McKnight's Long-Term Care News*. Accessed on: February 4, 2021. [Online]. Available: https://www.mcknights.com/blogs/guest-columns/nursing-home-quality-incentive-program-hardly/

[11] Ivorra, B., Ferrandez, M.R., Vela-Perez, M, and Ramos, A.M., (Apr. 30, 2020) "Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking into account the undetected infections. The case of China", *Elsevier Public Health Emergency Collection.* Accessed on: March 30, 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7190554/

[12] COVID Economics, "About Covid Economics", Centre for Economic Policy Research. Accessed on: Feb. 19, 2021. [Online]. Available: https://cepr.org/content/covid-economics-vetted-and-real-time-papers-0

[13] Prettenhofer, P., (Feb. 8, 2014) "Ordinary least squares in Python", *DataRoot*. Accessed on: Feb. 18, 2021. [Online]. Available: https://www.datarobot.com/blog/ordinary-least-squares-in-python/

AUTHORS

Dr. Omar Al-Azzam is an Associate Professor of Software Engineering in the Department of Computer Science and Information Technology (CSIT) at Saint Cloud State University (SCSU). Dr. Al-Azzam earned his BSc and MSc from Yarmouk University, Jordan and PhD from North Dakota State University (NDSU). Dr. Al-Azzam main research interests are big data analytics, bioinformatics and data mining.

Paul Court is a graduate student in the Professional Science Master of Software Engineering (PSMSE) program at Saint Cloud State University (SCSU) in the Department of Computer Science and Information Technology (CSIT). Mr. Court earned a MEd in Mathematics from the University of Minnesota and a BA in Mathematics from the University of Minnesota, Morris.

APPENDING SECURITY THEORIES TO PROJECTS IN UPPER-DIVISION CS COURSES

Vahab Pournaghshband[1] and Hassan Pournaghshband[2]

[1]Computer Science Department, University of San Francisco, USA
[2]Software Engineering & Game Development Department
Kennesaw State University, USA

ABSTRACT

Software systems have been under continued attacks by malicious entities, and in some cases, the consequences have been catastrophic. To tackle this pervasive problem, the academic world has significantly increased the offering of computer security-related courses during the past decade. In fact, offering these courses has become a standard part of the curriculum for many computing disciplines. While many proposals suggest adding this appealing topic into the non-security CS courses, many faculties do not entirely support the idea for a convincing reason. They rightfully claim that each one of these courses is already packed with concepts and materials developed toward that course, leaving not much room for other topics. In this study, we show how exposing students to security concepts can be incorporated into upper-division CS courses without increasing the normally required efforts needed by students as well as the instructor. We show how to develop a project of this nature that can be appended to an already existing course project. We have successfully employed our proposed approach in two of our core CS courses and present them in this paper as case studies.

KEYWORDS

Computer Science Education, Computer Security, Security Mindset.

1. INTRODUCTION

In recent decades, software has become a critical element in our lives. In parallel, the hackers have become increasingly capable of interrupting these technological connections by breaching software security. New threats are further emerging as computers become more embedded and play more intimately a big role into our environment and daily lives, as for the recent security vulnerabilities found in mobile medical devices [1]. In 2017, an attack was reported to a connected computer globally for every 39 seconds on average [2]. In the US, the number of data breaches increased by 900% in 2019 compared to 2005 [3, 4]. In the past year, over 500,000 Zoom account credentials were hacked and made available on the Dark Web [5]. In another instance, in the past year, a critical security flaw in WhatsApp was exploited, enabling hackers to install surveillance software on users' smartphones. This incident may have impacted WhatsApp's 1.5 billion users [6].

Computer security problems of this kind, and a variety of different ones, support the fact that it is critical for our students to gain the necessary skills and knowledge for handling them [7]. With this in mind, incorporating security into CS courses would have a positive effect on software security vulnerabilities, the most common cause of software security breaches. Consequently, a national-scale and critical unmet need exist for including fundamental principles of security in the

David C. Wyld et al. (Eds): CSTY, SIGI, MaVaS - 2021
pp. 27-37, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.112203

development of algorithms, software system implementation, and networked and mobile systems in primary undergraduate instruction. These days, undergraduate computer science students are not typically exposed to the rigorous constraints that security must impose in software development. The lack of experience and preparedness by students creates an increasing cost to the industry for not only training software engineers, also dealing with large risks associated with security weakness in their deployed products. Computing security is a central constraint in design, implementation, and usage of operating systems, database technology, networking, and other disciplines. Therefore, it must be integrated into each component of their curriculum. Among the reasons for the absence of security fundamentals in core undergraduate computing courses and curricula, one is the fact that there are relatively few faculties who bring a secure computing research background to a typical computer science department.

This paper discusses the challenges in teaching computer security and proposes a novel approach that can be effectively employed in teaching CS courses. This is a straightforward approach and is mainly based on introducing a security addition to an already existing project in non-security courses. It is important to note that, while the curriculum may include an elective upper-division security course, or have it incorporated in an introductory course, our intention is not to replace those courses, but rather to supplement them.

While basic security concepts can be taught in foundational programming courses (CS1), learning and absorbing the security concepts creditably requires advanced knowledge of computer science. This is why we believe that teaching them in upper-division courses complements what has been taught in introductory CS courses. Also, it is important to note that since security is not a standalone topic, its concept should be taught in context to be effective. These reasons are behind our strong belief that security should be taught in conjunction with the respective primary topics in non-security CS courses.

In this study, we show how exposing students to security concepts can be incorporated into upper-division CS courses without ever-increasing the efforts normally required by students and the actions needed by the instructor. We show how to develop a project of this nature that can be appended to an already existing course project. We have successfully employed our proposed approach in two of our core CS courses and present them in this paper as case studies.

The paper is organized as follows: Section 2 presents Related Work. Methodology is presented in Section 3 followed by Case Studies in Section 4. Section 5 concludes the paper.

2. RELATED WORK

There have been numerous works in literature focusing on computer security education [8-14]. In our prior work [8], we pragmatically examined major issues of teaching the security mindset in the early stage of programming by demonstrating how to teach the security mindset through a set of carefully crafted examples in lectures. Additionally, there have been studies emphasizing security concepts incorporation in early computer programming courses [8, 15, 16]. In this regard, educators have conducted various studies dealing with the challenge of how to incorporate security education into the undergraduate curriculum of the computer science or related discipline [17]. While some propose approaches for teaching a single course covering security concepts [17], some recommend effective track methods where a sequence of specialized security courses is offered [18]. As stated in [19], teaching security concepts and security mindset should be emphasized throughout the student's undergraduate program. Authors in [8, 20] and others have echoed this view, offering various teaching approaches.

3. METHODOLOGY

In this section, we present and discuss the methodology applied for our approach. We show how exposing students to security concepts can be incorporated into upper-division CS courses without increasing the required efforts normally needed by students as well as the instructor. We achieve this goal by developing an additional design phase that is security in nature, to an already existing course project. The project design phase is a combination of implementing both attack and defense mechanisms to the system implemented by the students in the original project. Since the project specification should be self-contained, it must include all the security terminologies, definitions, and preliminaries relevant to and required for the project. The main reason behind the project specification being self-contained is not requiring in-class discussions on security topics. This is because (1) the instructors might not be security experts, and (2), as the instructors rightfully claim, each one of these courses is already packed with concepts and materials developed toward that course, leaving not much room for other topics. The project specification also provides additional resources pertaining to security which enables students to conduct independent research in learning additional security concepts they may need in order to complete the project. Moreover, by appending a security phase to existing projects, our approach is easily adaptable to course projects currently employed in various upper-division courses.

The followings are the underlying rationale behind why appending a design project phase of security is naturally effective:

- It allows the students to make design choices. This way, they understand why one approach or algorithm is better than the other. For instance, through these projects, the students learn that AES encryption is more secure than DES encryption. As another example, they learn when to use public-key vs. private key encryption.
- Most curricula in introductory programming courses give students the impression that software is mostly correct, and security needs to be added to systems later, which overlooks the ‘secure first’ view. However, in our approach, we have a different objective by making the security design phase the last phase. We want the students to learn the principle of keeping security in mind from the beginning by understanding the complexity of adding it afterward.
- Design problems naturally lead to various solutions to the same problem based on inherently different design choices. Through end-of-semester in-class presentations, the students learn about other attack prevention or mitigation approaches presented by their classmates.

4. PROJECT DESIGN PHASE EXAMPLES: CASE STUDIES

In this section, we present four design projects that we have developed for the implementation of our proposed design project approach. These may be used as examples of how to add security components to already existing projects. We have recently incorporated these projects into our upper-division courses. We have done it for a senior-level operating system course and a junior-level computer network course. These courses were selected since they are presented to students who will be entering the industry workforce or transitioning to graduate computer science programs.

4.1. Operating Systems: Shell

4.1.1. Attack and defend your shell

The shell you implemented in the first phase of this project is a relatively functional shell compared to bash, but nevertheless has underlying security flaws. Identify these vulnerabilities and consequent potential attacks on the system running your shell that can be launched in your shell by a malicious user. Find them, identify to what extent these vulnerabilities can harm the system, and propose and/or implement a fix or mitigation technique to defend against such attacks. Some of these vulnerabilities are rather straightforward compared to subtle ones. For instance, use of printf and strcpy in your code already pose risks into your shell.

Extra credit: Exploit the vulnerabilities you have found in your lab1ab implementation to do something extremely harmful to the system running your shell (you may want to run it in a virtual machine.).

4.1.2. Process-Overload Attack

In an overload attack, a shared resource or service is overloaded with requests to such a point that it is unable to satisfy requests from other users. One of the simplest denial of service attacks is a process attack. In a process attack, one process or a user makes a computer unusable for other processes or users to run. The following program will probably paralyze or crash your system if your shell executes it, specially in root.

```
while(1)
      fork();
```

You will need to design a way to mitigate such attacks or perhaps not even let them happen. Find a way to secure the system using your shell from such attacks. (Hint: killing the rouge parent process will not solve the problem. You can try it).

4.1.3. Code Shell Injection

You will extend your implementation to introduce shell variables arguments to your ./timetrash, similar to how shell scripts handle arguments.

In that case, your ./timetrash (without -p or -t options) can have more arguments, which will be accessed by some of the commands listed in the script.sh.

Here is an example: `./timetrash script.sh myfile.txt cs`

Where the content of script.sh is: `cat $1 sort < $1 echo $2`

However, this feature can lead to security problems. In fact, a malicious user can inject and execute arbitrary commands. For instance, it can do some harm by running ./timetrash with the following arguments:

```
./timetrash script.sh "myfile.txt;rm -f /" cs
```

Or,

```
./timetrash script.sh "myfile.txt;mail evil@evilmail.com < /etc/shadow" cs
```

You are to implement variable shell arguments in a safe way to not let arbitrary commands to be executed.

4.2. Operating Systems: File Systems

4.2.1. Race Condition Attack

File systems are susceptible to race condition attacks on file systems when the following flaw in process behaviour is exploited: when a process performs a sequence of operations on a file, it assumes that the file does not change between any two successive operations. In this case, an attack can manifest itself by changing the file during the time window between two successive operations on it by a victim process. This temporal window is known as race window. There are two scenarios in which race condition in file systems may lead to potential damages: (a) the victim process operates on the changed file, leading to damage, or, (b) information is leaked/written from/into the file illegally if the victim's operations allow the attacker to get permissions on the file during the race window. Dictionary redirection attack, filelogger attack, and readfile attack are a few scenarios of race condition attacks on filesystems [21].

One way to prevent race condition is to use locks, which in this case could be called file locking. You are to design and implement file locking or any other approach that would prevent the attacks presented above in your implemented filesystem.

4.2.2. Filesystem-Level Encryption

As long as the operating system is running on a system without file encryption, access to the files will have to go through OS-controlled user authentication and access control lists. If an attacker gains physical access to the computer, however, this barrier can be easily circumvented. One way would be to remove the disk and put it in another computer with an OS installed that can read the filesystem, or simply reboot the computer from a boot CD containing an OS that is suitable to access the local filesystem.

The most widely accepted solution is to store the files encrypted on the physical media (disks, USB pen drives, tapes, CDs and so on). Recent operating systems have allowed users to store data encrypted on the filesystem (e.g., eCryptfs and EFS). Such systems provide both data and filename encryption on per-file basis. This is in contrast to full disk encryption where the entire partition or disk, in which the file system resides, is encrypted. In filesystem-level encryption, the file or folder is readable and writable only if the user provides the right password associated to that file or folder when they want to access it. Once implemented, this would enable files to be transparently encrypted to protect confidential data from attackers with physical access to the computer. By default, no files are encrypted, but encryption can be enabled by users on a per-file or per-directory basis.

You may need to add some information to the metadata of files in the underlying filesystem. This metadata describes the encryption for that particular file that is to be encrypted. It should provide both data and filename encryption on per-file basis.

You should decide which forms of encryption to support. In addition, you should address the security problem of keeping the key used for encryption in plain-text format in RAM, and how your approach mitigates this problem. In your approach of handling the encryption key in RAM problem, you should also consider the usability of your proposed security system.

4.2.3. Denial-of-Service Attacks on File systems

Denial-of-Service (DoS) attack is an attempt to make a machine unavailable to its intended users. Although the means to carry out, motives for, and targets of a DoS attack may vary, it generally consists of efforts to temporarily or indefinitely interrupt or suspend services on that system. One class of DoS attacks are those targeted for filesystems, specially that in Unix, files are more than just information storage. For instance, devices and sockets are also files.

You are going to propose a design an approach for all of the following particular filesystem DoS attacks and show why your design protects against such attacks.

4.2.3.1. Free space illusion attack

Open files that are unlinked continue to take up space until they are closed. The space that these files take up will not appear with the du or find commands, because they are not in the directory tree; however, they will nevertheless take up space, because they are in the filesystem.

For example:

```
int ifd;
char buf[8192];
ifd = open("./attack", O_WRITE|O_CREAT, 0777);
unlink("./attack");
while (1)
      write (ifd, buf, sizeof(buf));
```

Files created in this way can't be found with the ls or du commands because the files have no directory entries.

Hint: To recover from this situation and reclaim the space, you must kill the process that is holding the file open.

4.2.3.2. Deep directory attack

It is also possible to attack a system by building a tree structure that is made too deep to be deleted with the rm command. Such an attack could be caused by something like the following shell file:

```
$!/bin/ksh
$
$
Don't try this at home!
while mkdir anotherdir
do
      cd ./anotherdir
      cp /bin/cc fillitup
done
```

On some systems, rm -r cannot delete this tree structure because the directory tree overflows either the buffer limits used inside the rm program to represent filenames or the number of open directories allowed at one time.

4.2.3.3. Empty file attack

The UNIX filesystem uses inodes to store information about files. One way to make the disk unusable is to consume all of the free inodes on a disk, so no new files can be created. A person might inadvertently do this by creating thousands of empty files. This can be a perplexing problem to diagnose if you're not aware of the potential because the df command might show lots of available space, but attempts to create a file will result in a "no space" error. In general, each new file, directory, pipe, FIFO, or socket requires an inode on disk to describe it. If the supply of available inodes is exhausted, the system can't allocate a new file even if disk space is available.

You can tell how many inodes are free on a disk by issuing the df command with the -i option:

```
$ df -i
Filesystem    Inodes      IUsed       IFree     %IUse       Mounted on
/dev/sdb5    7569408     170084     7399324        3%       /
```

The output shows that this disk has lots of inodes available for new files.

Now if you run

```
$ touch empty_file
```

And run df -i again, you will see that it will increase the number of used inodes, even though it is just an empty file.

4.3. Operating Systems: Ramdisk

4.3.1. Encrypted Ramdisk

Recent operating systems have allowed users to store data encrypted on the filesystem. The filesystem is readable and writable only if the user provides the right password when they log in. Implement a software-encrypted ramdisk, where data is stored on the ramdisk in encrypted format. If a user opens the ramdisk normally, they should see encrypted gobbledegook. But if a user provides the right password at open time, then read operations on that open file should transparently decrypt the disk's data. Furthermore, if the user writes to the file, the data they write should be encrypted before it is sent to the ramdisk.

You will need to implement ramdisk-specific read and write operations for this design problem. (You probably need to change the osprd_blk_fops structure's read and write operations to point to your code.) You should decide which forms of encryption to support.

In addition, you should address the security problem of keeping the key used for encryption in plain-text format in RAM, and how your approach mitigates this problem. In your approach of handling the encryption key in RAM problem, you should also consider the usability of your proposed security system.

4.3.2. Partitioned Ramdisk

RAM was recently shown to be vulnerable to attacks exposing the totality of memory, including sensitive user data and encryption keys.

One way to mitigate data exposure is to divide the RAM into two partitions: protected partition to store sensitive data and a public partition to store everything else. This somewhat resembles partitioning a harddisk, with one difference: There is a hard boundary between the two partitions such that read and write accesses to the protected partition is only allowed by privileged processes.

This does not necessarily mean that the protected partition has to be encrypted. In fact, sometimes it is preferred not to be. One reason is that on-the-fly encryption and decryption often slows down access to memory (which defeats the purpose of using ramdisk for performance and speed in the first place). Also, some sensitive information should not be encrypted when in use anyways, such as encryption keys and passwords.

In your design, you should think about where and how in memory the information about the partitions you define is going to be kept. You should also not let the protected area of RAM to be swapped to disk, since this partition is designed to hold a variety of highly sensitive data, which will be exposed if it is swapped to harddisk. Even by adding such security, some form of physical attack, called cold boot attack, has proven to expose data in the memory even when the system is powered off.

By incorporating your design can we stop unprivileged processes from reading the content of the protected memory using “memory viewers”? Is it possible to protect against cold boot attack?

(Hint: One design to mitigate this is to introduce privileged read and write system calls)

4.4. Computer Network: Peer-to-Peer File Sharing

4.4.1. Authentication and Authorization

Authentication is the process carried out by an entity to confirm the identity of another entity or to confirm that a data is indeed from whom it claims to be. Passwords, digital signatures or message authentication codes are the common techniques of authentication. Authorization on the other hand, is the process of granting to the users some privileges on the access to a set of resources according to what is permitted to them. The most popular techniques to enforce authorization are to maintain access control lists (ACL) listing the access rights of entity.

The peer is happy to serve other peers any data contained in its current directory. Fancier programs, such as real web servers, allow users to specify which files in a directory can be served. For example, this syntax tells Apache to refuse to serve the osppeer.c file:

```
<Files "osppeer.c"> Order allow,deny Deny from all </Files>
```

This syntax would usually go in a file called .htaccess, in the directory containing osppeer.c.

Design access control syntax for our peers. Will you support Apache-style .htaccess files, or something else? What type of syntax will you support? For full credit, you should design a very flexible access control syntax. Consider such issues as limiting access for some files to limited sets of peers, defined based on (say) network address; symbolic links; and so forth.

In addition to authorization, your design should also authenticate the peer via some ACL mechanism you will design, so that you serve the file only if the peer is authenticated and is authorized to download the file from you.

4.4.2. Transmitting Encrypted Files

Extend our current design to allow peers to send encrypted data. You should consider three types of encryption. In increasing order of safety:

1. Hiding file contents from network snoopers.
2. Hiding file contents from unauthorized peers. I.e., if a peer does not know the right key(s), then the peer will not be able to understand a download file.
3. Hiding the existence of a file from unauthorized peers. I.e., if peer 1 does not know the right key(s), then peer 1 cannot tell which files peer 2 has made available. (Perhaps peer 1 will be able to tell that peer 2 has registered 5 files with weird scrambled names, but peer 1 cannot tell what those files' true names are, and peer 1 cannot download their data -- encrypted or not -- from peer 2.)

For each of the items mentioned above, you should address what key management mechanism you put in place in this peer-to-peer network. In doing so, your design should handle scenarios such as where the key is compromised or a scenario where peer 1 would like to revoke peer 2's access to a particular file it is sharing, while peer 3 and peer 4 would still be able to access the file, as they originally have. How would your encryption system differ in such cases?

Extra credit: Confidentiality is reached when data is protected from unauthorized disclosures, whereas integrity is means that data is safe from unauthorized modifications. Encryption is a powerful guarantee of confidentiality, while data integrity can be achieved using digital signatures and message authentication codes. Replay attacks involve a malicious user injecting old data on the system. In order to guarantee freshness, timestamps or nonces can be used, but these tools require a certain degree of synchronization between the entities. How can you extend your cryptographic protocol to ensure authenticity of content? (Hint: one way to achieve that is through signed public-key certificates by trusted publishers)

5. CONCLUSIONS

While software has become a crucial element in our lives in recent decades, and consequently, software systems have been under continued attacks by malicious entities. As a result, in recent years, the offering of computer security courses in universities has increased. While many proposals suggest adding this appealing topic into the non-security CS courses, there are several challenges in implementing them effectively. We addressed those challenges by presenting a novel approach to show how to develop a project of security nature that can be appended to an already existing course project. We employed our proposed approach in two of our Operating Systems and Computer Networks courses and present the details as case studies.

REFERENCES

[1] Vahab Pournaghshband, Majid Sarrafzadeh, and Peter Reiher (2012) “Securing legacy mobile medical devices”, *In Proc. of 3rd International Conference on Wireless Mobile Communication and Healthcare* (MobiHealth ’12). Springer.

[2] The Security Magazine (2017) “Hackers Attack Every 39 Seconds.”, https://www.securitymagazine.com/articles/87787-hackers-attack-every-39-seconds.

[3] DataStealth (2020) “110 Must-Know Cybersecurity Statistics for 2020”, https://www.datex.ca/blog/110-must-know-cybersecurity-statistics-for-2020.

[4] Statista (2021) “Number of compromised data records in selected data breaches as of January 2021”, https://www.statista.com/statistics/290525/cyber-crime-biggestonline-data-breaches-worldwide/.

[5] Forbes.com (2020) "Hacked Zoom Accounts Given Away For Free On The Dark Web", https://www.forbes.com/sites/leemathews/2020/04/13/500000-hackedzoom-accounts-given-away-for-free-on-the-dark-web/#6e67407a58c5.

[6] BBC Technology (2019) "WhatsApp discovers targeted surveillance attack", https://www.bbc.com/news/technology-48262681.

[7] Computing Curricula (2017) "Curriculum guidelines for post-secondary degree programs in cybersecurity" *IEEE Computer Society Association for Computing Machinery,* https://doi.org/10.1145/3184594

[8] Vahab Pournaghshband (2013) "Teaching the security mindset to CS1 students", *In Proceeding of the 44th ACM technical symposium on computer science education* (Denver, Colorado, USA) (SIGCSE '13). ACM, 6. http://doi.acm.org/10.1145/2445196.2445299

[9] Debarati Basu, Harinni K. Kumar, Vinod K. Lohani, N. Dwight Barnette, Godmar Back, Dave McPherson, Calvin J. Ribbens, and Paul E. Plassmann (2020) "Integration and Evaluation of Spiral Theory Based Cybersecurity Modules into Core Computer Science and Engineering Courses" *In ACM SIGCSE '20* Association for Computing Machinery, 9–15.

[10] Trajce Dimkov,Wolter Pieters, and Pieter Hartel (2011) "Training students to steal: a practical assignment in computer security education", *In Proceedings of the 42nd ACM technical symposium on computer science education*. ACM, 21–26.

[11] Tadayoshi Kohno and Brian D Johnson (2011) "Science fiction prototyping and security education: cultivating contextual and societal thinking in computer security education and beyond", *In Proceedings of the 42nd ACM technical symposium on Computer science education*. ACM, 9–14.

[12] Peter Peterson and Peter L Reiher (2010) "Security Exercises for the Online Classroom with DETER", *In Proceedings of the 3rd international conference on Cyber security experimentation and test* (CSET'10).

[13] Valdemar Šv.benský, Jan Vykopal, and Pavel Čeleda (2020) "What Are Cybersecurity Education Papers About? A Systematic Literature Review of SIGCSE and ITiCSE Conferences", *In Proceedings of the 51st ACM Technical Symposium on Computer Science Education* (Portland, OR, USA) (SIGCSE '20). 2–8.

[14] Bernard Yett, Nicole Hutchins, Gordon Stein, Hamid Zare, Caitlin Snyder, Gautam Biswas, Mary Metelko, and Ákos Lédeczi (2020) "A Hands-On Cybersecurity Curriculum Using a Robotics Platform" *In ACM SIGCSE'20*. Association for Computing Machinery, 1040–1046.

[15] Vahab Pournaghshband and Hassan Pournaghshband (2021) "Teaching Security Notions in Entry-Level Programming Courses", *In Proc.* of *IEEE International Conference on Teaching, Assessment, and Learning for Engineering* (IEEE TALE '21). IEEE.

[16] K. Nance (2009) "Teach Them When They Aren't Looking: Introducing Security in CS1" *In IEEE Security Privacy* 7, 5 (Sept 2009), 53–55. https://doi.org/10.1109/MSP.2009.139

[17] Luiz Felipe Perrone, Maurice Aburdene, and Xiannong Meng (2005) "Approaches to undergraduate instruction in computer security", *In Proceedings of the American Society for Engineering Education Annual Conference and Exhibition,* ASEE.

[18] S. Azadegan, M. Lavine, M. O'Leary, A. Wijesinha, and M. Zimand (2003) "An Undergraduate Track in Computer Security", *SIGCSE Bull.* 35, 3 (June 2003), 207–210. https://doi.org/10.1145/961290.961568

[19] Ambareen Siraj, Nigamanth Sridhar, John A. Drew Hamilton, Latifur Khan, Siddharth Kaza, Maanak Gupta, and Sudip Mittal (2021) "Is There a Security Mindset and Can It Be Taught?" *In Proceedings of the Eleventh ACM Conference on Data and Application Security and Privacy* (CODASPY'21). Association for Computing Machinery, New York, NY, USA, 335–336. https://doi.org/10.1145/3422337.3450358

[20] Kara Nance and Brian Hay (2009) "Computer Security-focused Programming Assignments in Foundational CS Courses" *In Proceedings of the 13th Colloquium for Information Systems Security Education*.

[21] Prem Uppuluri, Uday Joshi, and Arnab Ray. (2005) "Preventing race condition attacks on file-systems", *In Proc. of the ACM symposium on Applied computing* (SAC '05). Association for Computing Machinery, New York, NY, USA, 346–353.

AUTHORS

Vahab Pournaghshband is an Associate Professor in the Computer Science department at University of San Francisco. He received his Ph.D. in Computer Science at University of California, Los Angeles (UCLA) in 2014. He received his M.Sc. in Computer Science from University of California, Berkeley. Also from UC Berkeley, he received his B.Sc. in Electrical Engineering and Computer Science. Dr. Pournaghshband was a recipient of Fulbright US Scholar Award and has published numerous papers and has made numerous national and international professional presentations at different conferences. His current area of research are Computer Networks, Computer Security, and Computing Education.

Hassan Pournaghshband, recipient of the Kennesaw State University 2021 Outstanding Teaching Award, is a professor of software engineering in the department of software engineering and game development at KSU. Before joining KSU he taught at City University of New York, Hofstra University, University of Tehran, and University of Missouri. He earned his bachelor's degree in Economics from University of Tehran, his Masters of Computer Science from Northwestern University, and his Ph.D. in Computer Science from University of Oklahoma. He has published numerous papers and has made numerous national and international professional presentations at different conferences. His current areas of research are Software Architecture, Software Project management, Software QualityAssurance and testing, and Database Management Systems.

EMOTIONS IN VIRTUAL REALITY

Darlene Barker and Haim Levkowitz

Department of Computer Science,
University of Massachusetts Lowell, Lowell, MA, USA

ABSTRACT

One of the first senses we learn about at birth is touch, and the one sense that can deepen our experience of many situations is touch. In this paper we propose the use of emotions including touch within virtual reality (VR) to create a simulated closeness that currently can only be achieved with in-person interactions and communications. With the simulation of nonverbal cues, we can enhance a conversation or interaction in VR. Using haptic devices to deliver the simulation of touch between users via sensors and machine learning for emotion recognition based on data collected; all working towards simulated closeness in communication despite distance or being in VR. We present a direction for further research on how to simulate in-person communication within VR with the use of emotion recognition and touch to achieve a close-to-real interaction.

KEYWORDS

Virtual reality, emotion recognition, deep learning, in-personal communication simulation, nonverbal communication.

1. INTRODUCTION

To make an impact on social interaction within virtual reality (VR), we need to consider the effect of emotions in our in-person communications and their expression within VR [12]. The discussion "and theorizing about what makes up a virtual reality or virtual experience" has been going on since the 1960s and even further back when it was a dream or a work of fiction [6]. Now that virtual reality is a reality and VR headsets have become fixtures in a growing number of households, we need to utilize this technology to work towards a more impactful social interaction and establishing the connectedness we crave as humans [4].

One of the drawbacks of current forms of long-distance communication methods is that they lack the physical, nonverbal contact and expression that "form an important part of information exchange in our everyday interpersonal experiences" [2]. We rely on "realistic sensory simulations" for visual, tactile, and audio cues [9]. What we are proposing is simulating this in VR with the use of emotions. There are four pieces to the puzzle of making a long-distance interaction feel intimate within VR—emotions, voice, facial expressions, and touch. We express our emotions in day-to-day communications using the other three pieces. To do the same in VR, we need a novel way of implementing touch. Now, the goal becomes making the experience as close to the real world as possible, while also keeping it interesting and achievable.

When a person speaks, a lot can be interpreted from the acoustic properties of what they said; these can be features that are recorded and used to help interpret their emotional state. This can include the pitch at which they spoke and how loud they were. Facial expressions are changes in facial muscles that can show emotions, including the basic ones, such as, happy, sad, surprise,

David C. Wyld et al. (Eds): CSTY, SIGI, MaVaS - 2021
pp. 39-46, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.112204

fear, anger. This happens when "a series of biological events follow that produce changes in a person" [1]. You are happy to see the other person on the line and you smile with your whole face. That is a favorable reaction but then let us say, you are receiving disappointing news: your facial reaction will be quite different.

The sense of touch is one of those senses that a newborn can experience for the first time right out of the womb with their mother. It is "the first of our senses to develop, and it provides us with our most fundamental means of contact with the external world" making this an important part of emotion and the sense of presence in the real world [2]. So, in trying to recreate this in VR, touch becomes a critical part of the formula of an immersive experience. Not only does touch affect our emotions, but it also affects how we act, make connections to, and interact with others.

During in-person interactions, we use verbal and nonverbal body language cues to communicate. Long-distance communication like text messages, phone calls, video calls, and handwritten letters will only offer part of the whole experience. Within a virtual environment (VE), we can make long-distance communications closer to in-person experiences and add a level of intimacy using emotion recognition and touch utilization. We propose "smart clothes" along with head mounted displays to create such a virtual environment [3]. At this time, we can only provide feedback on the controllers for hand holding and headsets that are currently available. However further research is needed for the entire body. Being able to see the face of the person and their body movements, along with the added feature of touch, creates a greater degree of intimacy in long-distance communication.

"Virtual reality (VR) allows for a close approximation of the real world," which is what we are using to create an intimate long-distance communication within a "controllable approximation of the real, physical world" [8]. We propose to use VR for long-distance communications where the avatars of the parties reflect their emotions, and each party can touch each other. When they touch, they will receive a signal and have an emotional response that will be visible to the party initiating the touch. The emotions create a sense of presence that helps with the immersive element of this type of communication.

2. OUR APPROACH

We are proposing the use of passive sensors to capture physiological changes during interaction in a virtual environment as well as sensors that capture deliberate actions on either party. So, one party reaches out to touch the arm of the other party; the sensors would collect data and that data would assist in the simulation of the feeling of being touched by the other party. This would happen in real time for both parties. When you touch a person's arm, you feel the arm and at same time, you also feel a response. The other party feels this touch and may or may not respond actively. So, this is where the passive sensors can be of assistance. An example could be feeling a cold object on your arm. One could respond with that was not cold, but the arm shivers a tiny bit; and delivers a response that would have been seen if you were next to the person. We are aiming towards a more real simulation of touch and emotional interaction in VR.

The idea is to create a complete environment "or world" for the users to experience a close interaction with each other even if they are separated by distance [6]. We are proposing "a unique interaction space where users can perceive a different reality" than their own, one in which distance does not affect the interaction as it does in the real world [6]. The idea is having two users interact in real time in VR in a computer-generated VE. Instead of using objects in the world to stimulate the user's emotions and affect their sense of presence, we are proposing the capture of their emotions from an evaluation of their speech, facial expression, physical changes in their body, body language, and how they are interacting with the world [7].

We propose doing this using smart garments and other VR wearable devices that can read input and transfer signals to and from the wearer of the garment or device [3], [11], [21]. This is one of the goals of this research. While VR can be seen as having a need "to create a sense of 'connection' to evoke feelings of presence and enhance user engagement", we are aiming for a simulation of the physical experience [23]. Current VR experiences "touch illusion relying on visuals and sound is thought to increase immersion", we are aiming at simulating touch which will bring us closer to the mark in recreating emotions in VR [21], [23].

One may ask, why touch? Well, it is the first sense that we develop as an infant. Often, our first experience in the world is the touch of the mother. Touching an object can make the experience more real than simply looking at it; it gives the connectedness that we want in VR which will immerse the user in the virtual world. "The sense of touch is a collection of several sensations, encompassing pressure, pain and temperature, and touch receptors are stimulated by a combination of mechanical, chemical and thermal energy" [11]. By capturing those sensations along with the physiological changes in the body, we are aiming to add touch to the experience in VR "to deliver something that can approximate to realistic interpersonal touch" [2].

Much of the data collection will need to be evaluated in context since data will be collected from multiple sensors because of many changes that occur in our bodies for a single emotion to be expressed [17]. To then recreate an emotion such as happiness; the eyes are wide open, there may be some excitement, the palms a little sweaty, the face may be smiling, the back may be straight, and the upper body may be open and confident, and the voice may be high pitched. All of these depend on the baseline of the users.

Data collection could include the use of cameras within the headset and outside to capture facial expressions potentially including eye-tracking, movements of the body, and any body language cues [16], [18]. Other data collection could include an EEG sensor to capture brain waves, and other sensors collecting the temperature of the skin, galvanic skin response (GSR), speech, eye-tracking, heart rate [13], [18], [19], [20]. Speech can be captured and transmitted but there can be patterns in the way things are being said that could give a clearer picture of the emotional state of the spoken part. We will use techniques from machine learning to process this data that we collect on both parties and recreate each party visually (avatar) as well as the feelings associated with touch [18], [22]. The idea will be to produce an avatar that matches the emotional state of the person the data was collected from without any input from the person being captured.

One of the techniques that is being considered is a LSTM based deep learning network to assist in detecting the emotions, as well as the effect of what one is hearing in the virtual environment, such as music, or other sounds including the speech of the other party [14], [15].

Table 1. Communication options and senses used

Long-distance communication Options	**Senses**		
	Hearing	***Sight***	***Touch***
Phone calls	Yes		
Emails			
Video chat apps	Yes	Yes	
Handwritten letters			
VR Chat	Yes	Yes	
Virtual Touch	Yes	Yes	Yes

In Table 1 we see communications options and senses that we utilize. Of the senses being covered in communication, hearing and sight are the ones covered most, and we are proposing to improve communications by adding touch to the mix through indirect observation via the use of passive sensors. We will create a VE setting and use emotions to help immerse the user into the world and be able to physically feel feedback beyond the feedback of a vibrating controller. We propose the use of emotion tracking which is used to interpret what both parties are feeling and deliver that to each other along with simulating touch to enhance the experience and what is being communicated.

3. RELATED WORK

The closest option to what we are trying to accomplish is the use of virtual interpersonal touch with expression and recognition of emotions using force-feedback haptic. Emotions can be calibrated from facial expression data detected from facial anchor points using machine learning, digitized audio streams, and data collected from virtual interpersonal touch [1]. Emotion data can be in the form of a 2D scale of the hedonic valences and intensity. Facial expression shows emotion in the form of the facial muscles and change of the heart rate or blood pressure. Voice can be gotten from digital audio streams.

Touch tends to increase trust in face-to-face communication, and it usually adds sincerity, establishes trust, weight, urgency, marks significance, or with greetings (shake hands). It is used in one-to-one interactions, seen as inefficient and inconveniently requiring proximity [1]. It was also found that “overly intimate or socially inappropriate for many interactions”, “little is known about the psychological effects of haptic communication”, “can create transformed haptic communication scenarios that are not possible in the physical world”, and “communication of emotion in virtual reality does not necessarily have to copy the real world; instead, it can be abstracted” [1].

This study was found to not be as accurate as interpersonal handshakes and inefficient because the experiment required the subject to be connected to multiple devices. Touch was found to be too intimate if not socially inappropriate in some cases, but at the same time they were able “to convey emotions via a short interaction such as a handshake” [1].

For our approach, more than a handshake is necessary. We are trying to convey intimacy across distance. The idea is to be able to touch the hand of the other person as you would while having dinner or touch the arm or shoulder or some other show of affection.

Another paper by Gallace and Spence, talks about interpersonal touch and how it can be used to communicate emotion, improve our wellbeing, to share our feelings with others, and to enhance the meaning of other forms of verbal and non-verbal communication [2]. They also mention the “recent growth of studies investigating the addition of tactile sensations to long-distance communication technologies” in the form of “mediated or ‘virtual’ touch” [2]. For our approach touch is an integral part that is missing from current virtual and long-distance communication options.

Another paper that included experiments (shown below) with the use of facial expressions and the use of touch, and in this case, a tactile glove is used for participants to be able to shake hands. This study found that “not only does touch affect emotional processes, but the processing of touch may also be” altered by other factors and may not be a positive emotion [4]. This does not cover what we are proposing in the executing of touch but does cover capturing other emotions.

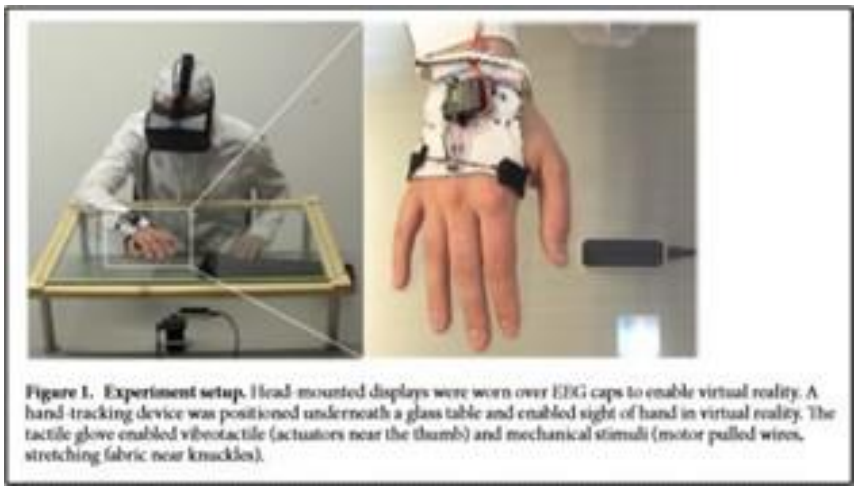

Figure 1. Experiment setup. Head-mounted displays were worn over EEG caps to enable virtual reality. A hand-tracking device was positioned underneath a glass table and enabled sight of hand in virtual reality. The tactile glove enabled vibrotactile (actuators near the thumb) and mechanical stimuli (motor pulled wires, stretching fabric near knuckles).

Figure 1. taken from Ravaja, N. et al [4]

We propose to do so using senses versus visual manipulation. As humans we are aware of our surroundings based on what we see, touch, smell, and feel. With the ability to touch and feel feedback from the virtual environment, we would have a more realistic sense of presence.

Another paper looked at altering the mood of the participants using touch and smell. The touch component was objects in the world, such as grass. There was a change to the person in the real world that affected their mood in the virtual world. They found that "the more senses stimulated, the higher the sense of presence" for the participants in the study [7]. For our approach, we are looking at creating the stimulation within the virtual environment.

In [10], Hinkle proposed a machine learning approach to extracting physical measurements "to determine emotional state of human test subjects through the non-invasive collection and analysis of physiological data during subject interaction virtual reality or other immersive audio/visual environments". Their approach proposes "a circumplex model of affect, more widely known as the arousal-valence model" using "image processing and machine learning techniques" where they were about to "classify emotion based on facial images and voice recordings" [10]. Our approach is to also collect the data with non-invasive means as well as machine learning to extract the features.

In [11], the authors discuss a current technology used for enhanced touch, a "skin-integrated technology that applies pressure, vibration or motion to the user, enabling communication between the user and a machine for virtual and augmented reality," which "have some drawbacks. For instance, each actuator is driven by a set current of about five milliamps, which is relatively high compared with that found in other consumer digital electronics" [11]. Each actuator is "equipped with embedded sensors that detect and encode the strength, duration and location of the touch, together with skin warmth and heart rate of the sender" [11]. This is the closest to what we are proposing but further research is needed.

4. POTENTIAL IMPACT ON SOCIETY

This is an opportunity to make social interactions in VR more real since the addition of touch and the integration of emotions in the whole experience will add a level of accountability. This could make for more honest interaction in VR. This creates a new and viable means to maintain

personal relationships when used for communication between parties who are separated by distance for extended periods of time.

Other areas that this could be applied to include gaming, prosthetics, and stress management. In gaming interactions in VR, it could add to the immersive nature of the technology. This could potentially be of a significant help in "prosthetic feedback" [11]. In the case where VR is being used to teach emotion recognition, a way to integrate touch could be impactful to the immersive requirement of the experience. This can be a tool for stress reduction since a "tender caress from a caregiver or spouse reduces stress and promotes emotional well-being" [5]. The stress of separation from loved ones can be reduced with the use of touch and emotion recognition in VR.

5. FUTURE WORK

The perceptual immersive nature of VR with the use of voice, face, and touch to execute a long-distance communication experience could give the illusion of being a step closer towards a nonverbal intimacy and a certain degree of emotional warmth. The reasoning for using touch in VR is establishing trust in interpersonal interactions gives the illusion of the closeness that is a necessary part of maintaining connections when separated by great distances.

VR has been around for some time and now it is becoming a fixture in a growing number of households; the next logical step would be to have it evolve into a tool that can help the human experience. VR can help communication across continents and great distances, provided the necessary power and Wi-Fi is available, and assist with maintaining the closeness that is lacking in current video, telephone, email, or text communication—which is all that is afforded in the current market.

Its not just VR but simulating touch via electronic technology can have many applications—say a visually impaired person can feel the skin of a loved one across the globe, as in:

- In our current normal, holding the hand of a loved one in a situation where physical contact cannot always happen, as in a medical facility, during this time of quarantine.
- Being able to be there for a loved one despite distance or circumstances, as in soldiers being deployed or a parent, or loved one away for work being able to experience important moments in our lives, such as, birth of a child, or some other special occasion. Being able to hold your partner, parent, or child's hand in those important moments despite being separated because of the pandemic or distance.

Granted, this approach requires further research on technology to make touch detection and transfer between two parties available to the consumer market, and integration with current VR head mounted devices. With all the above, this is still a novel idea to further the human experience.

6. CONCLUSION

Our goal is to bring people together despite distance in such a way that replicates and even rivals in-person communication. It could be used in video games, to enrich personal relationships, virtual healthcare services, virtual education, retail, therapeutic services, and potentially in other situations where physical interaction is not possible.

REFERENCES

[1] Bailenson, J., Yee, N., Brave, S., Merget, D., Koslow, D., "Virtual Interpersonal Touch: Expressing and Recognizing Emotions Through Haptic Devices," Human-Computer Interaction, 2007, Volume 22, pp. 325-353.

[2] Gallace, A., Spence, C., "The science of interpersonal touch: An overview", Neuroscience and Biobehavioral Reviews, 34, 2010, pp. 246-259.

[3] Makki, I., Alhalabi, W., Adham, R., "Using Emotion Analysis to Define Human Factors of Virtual Reality Wearables", ScienceDirect, Procedia Computer Science 165, 2019, pp. 154-164.

[4] Ravaja, N., Harjunen, V., Ahmed, I., Jacucci, G., Spape, M., "Feeling Touched: Emotional Modulation of Somatosensory Potentials to Interpersonal Touch", Scientific Reports 7:40504, 2017, pp. 1-11.

[5] Harjunen, V. J., Spape, M., Ahmed, I., Jacucci, G., Ravaja, N., "Individual differences in affective touch: Behavioral inhibition and gender define how an interpersonal touch is perceived", Personality and Individual Differences, 107, 2017, pp. 88-95.

[6] Cisners, Angela, Maravilla, Morelia, Murray, Brian, Scretching, Desiree, Stoddard, Alvie, Redmiles, Elissa M., "Defining Virtual Reality: Insights from Research and Practice", iConference 2019 Proceedings, 2019.

[7] Serrano, B., Banos, R., Botella, C., "Virtual reality and stimulation of touch and smell for inducing relaxation: A randomized controlled trial", Computers in Human Behavior 55, 2016, pp. 1-8.

[8] van der Ham, I., Klaassen, F., Van Schie, K., Cuperus, A., "Elapsed time estimates in virtual reality and the physical world: The role of arousal and emotional valence", Computers in Human Behavior 94, 2019, pp. 77-81.

[9] Lecuyer, Anatole, "Playing with Senses in VR: Alternate Perceptions Combining Vision and Touch", Spatial Interfaces 0272-1716, 2017, pp. 20-26.

[10] Hinkle, Lee B., Roudposhti, Kamrad K., Metsis, "Vangelis, Physiological Measurement for Emotion Recognition in Virtual Reality", International Conference on Data Intelligence and Security, 2019, pp. 136-143.

[11] Tao, Xiao-ming, "Virtual and augmented reality enhanced by touch", Nature, Vol. 575, 2019, pp. 453-454.

[12] Barker, Darlene, Levkowitz, Haim, "Tutorial 5: Emotions in Virtual Reality", IEEE VR 2021 Conference, https://ieeevr.org/2021/program/tutorials/#T5

[13] Yang, Xiuleng, Emma McCoy, Esther Anaya-Boig, Ione Avila-Palencia, Christian Brand, Glòria Carrasco-Turigas, Evi Dons, et al, "The Effects of Traveling in Different Transport Modes on Galvanic Skin Response (GSR) as a Measure of Stress: An Observational Study", Environment International 156 (November), 2021, 106764.

[14] Sakalle, Aditi, Pradeep Tomar, Harshit Bhardwaj, Divya Acharya, and Arpit Bhardwaj, "A LSTM Based Deep Learning Network for Recognizing Emotions Using Wireless Brainwave Driven System", Expert Systems with Applications 173 (July), 2020, 114516.

[15] S. Sheykhivand, Z. Mousavi, T. Y. Rezaii and A. Farzamnia, "Recognizing Emotions Evoked by Music Using CNN-LSTM Networks on EEG Signals", IEEE Access, Vol. 8, 2020, pp. 139332-139345.

[16] Geraets, C.N.W., S. Klein Tuente, B.P. Lestestuiver, M. van Beilen, S.A. Nijman, J.B.C. Marsman, and W. Veling, "Virtual Reality Facial Emotion Recognition in Social Environments: An Eye-Tracking Study", Internet Interventions 25 (September), 2021, 100432.

[17] B. Meuleman and D. Rudrauf, "Induction and Profiling of Strong Multi-Componential Emotions in Virtual Reality", IEEE Transactions on Affective Computing, Vol. 12, no. 1, Jan.-March 2021, pp. 189-202.

[18] Magdin, Martin, Zoltán Balogh, Jaroslav Reichel, Jan Francisti, Stefan Koprda, and Molnár György, "Automatic Detection and Classification of Emotional States in Virtual Reality and Standard Environments (LCD): Comparing Valence and Arousal of Induced Emotions", Virtual Reality 25, No. 4 (December 1, 2021): 1029.

[19] Marín-Morales, Javier, Juan Luis Higuera-Trujillo, Jaime Guixeres, Carmen Llinares, Mariano Alcañiz, and Gaetano Valenza, "Heart Rate Variability Analysis for the Assessment of Immersive Emotional Arousal Using Virtual Reality: Comparing Real and Virtual Scenarios." PLoS ONE 16, No. 7 (July 2021): 1–16.

[20] Gilpin, Gina, James Gain, and Gosia Lipinska, “The Physiological Signature of Sadness: A Comparison between Text, Film and Virtual Reality”, Brain and Cognition 152 (August 1, 2021).
[21] Yu, X., Xie, Z., Yu, Y. et al, “Skin-integrated wireless haptic interfaces for virtual and augmented reality”, Nature 575, 473–479 (2019).
[22] Jianhua Zhang, Zhong Yin, Peng Chen, Stefano Nichele, “Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review”, Information Fusion, Volume 59, 2020, pp 103-126.
[23] Price, Sara & Jewitt, Carey & Yiannoutsou, Nikoleta, “Conceptualising touch in VR”. Virtual Reality, 25, 2021, 10.1007/s10055-020-00494-y.

AUTHORS

Darlene Barker is a PhD Candidate in the Computer Science Department at the University of Massachusetts Lowell, in Lowell, MA, USA. She completed her master’s degree in Computer Science from University of Massachusetts. She also completed her Bachelor of Science majoring in Computer Science from Rivier University in Nashua, NH. Her professional career experience spanning several years and disciplines from systems to software working for companies—Compaq Computer Corporation, Teradyne, Inc, and Oracle Corporation. She is the founder of an independent video game studio, DB Attic Studios, LLC, in Bedford, NH where she makes video games with her children from middle school through high school graduates.

Haim Levkowitz is the Chair of the Computer Science Department at the University of Massachusetts Lowell, in Lowell, MA, USA, where he has been a faculty member since 1989. He was a twice-recipient of a US Fulbright Scholar Award to Brazil (August – December 2012 and August 2004 – January 2005). He was a Visiting Professor at ICMC — Instituto de Ciencias Matematicas e de Computacao (The Institute of Mathematics and Computer Sciences)—at the University of Sao Paul, Sao Carlos – SP, Brazil (August 2004 - August 2005; August 2012 to August 2013). He co-founded and was Co-Director of the Institute for Visualization and Perception Research (through 2012). He is a world-renowned authority on visualization, perception, color, and their application in data mining and information retrieval. He is the author of “Color Theory and Modeling for Computer Graphics, Visualization, and Multimedia Applications” (Springer 1997) and co-editor of “Perceptual Issues in Visualization” (Springer 1995), as well as many papers in these subjects. He is also co-author/co-editor of "Writing Scientific Papers in English Successfully: Your Complete Roadmap," (E. Schuster, H. Levkowitz, and O.N. Oliveira Jr., eds., Paperback: ISBN: 978-8588533974; Kindle: ISBN: 8588533979, available now on Amazon.com: http://www.amazon.com/Writing-Scientific-Papers-English-Successfully/dp/8588533979). He has more than 48 years-experience teaching and lecturing, and has taught many tutorials and short courses, in addition to regular academic courses. In addition to his academic career, Professor Levkowitz has had an active entrepreneurial career as Founder or Co-Founder, Chief Technology Officer, Scientific and Strategic Advisor, Director, and venture investor at several high-tech start-ups.

.

A NOVEL ATTENTION-BASED NETWORK FOR FAST SALIENT OBJECT DETECTION

Bin Zhang, Yang Wu, Xiaojing Zhang and Ming Ma

College of Computer Science and Engineering,
Inner Mongolian University, China

ABSTRACT

In the current salient object detection network, the most popular method is using U-shape structure. However, the massive number of parameters leads to more consumption of computing and storage resources which are not feasible to deploy on the limited memory device. Some others shallow layer network will not maintain the same accuracy compared with U-shape structure and the deep network structure with more parameters will not converge to a global minimum loss with great speed. To overcome all of these disadvantages, we proposed a new deep convolution network architecture with three contributions: (1) using smaller convolution neural networks (CNNs) to compress the model in our improved salient object features compression and reinforcement extraction module (ISFCREM) to reduce parameters of the model. (2) introducing channel attention mechanism in ISFCREM to weigh different channels for improving the ability of feature representation. (3) applying a new optimizer to accumulate the long-term gradient information during training to adaptively tune the learning rate. The results demonstrate that the proposed method can compress the model to 1/3 of the original size nearly without losing the accuracy and converging faster and more smoothly on six widely used datasets of salient object detection compared with the others models. Our code is published in https://gitee.com/binzhangbinzhangbin/code- a-novel-attention-based-network-for-fast-salient-object-detection.git

KEYWORDS

Salient Object Detection, Optimization Strategy, Deep Learning, Model Compression, Vision Attention.

1. INTRODUCTION

Salient object detection plays an important role in many computer vision tasks, such as computer vision tracking [1], content aware image processing [2], medical segmentation, robot vision navigation [3] and so on. Most of traditional detection methods [4,5,6,7,8,9,10,11] extract handle-crafted features to capture low-level visual features such as color, intensity and orientation. Traditional models are also dependent on the manually designed regional salient descriptors. However, low-level semantic information which means local contrast operator has a limited spatial neighborhood and large high-level feature information can be easily missed out. With the advent of convolution neural network, it can not only capture the low-level semantic information but also extract the high-level semantic information, which greatly improves the accuracy of significance by fusing the context details features of multi-scale space.

David C. Wyld et al. (Eds): CSTY, SIGI, MaVaS - 2021
pp. 47-58, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.112205

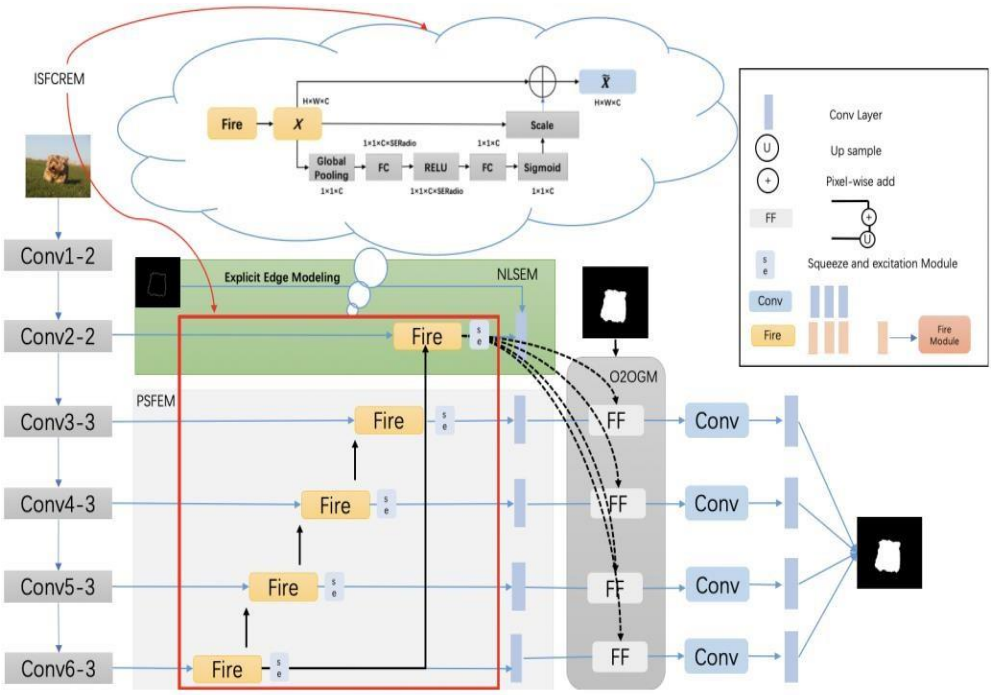

Figure 1. The pipeline of the proposed approach

Figure 1. The pipeline of the proposed approach. Based on the Edge Guidance Network we replace the convolution layers of the decoder of U-shape structure with fire module. Follows we use squeeze-and-excitation module to focus more attention on different channel. PSFEM: progressive salient object features extraction module. NLSEM: non-local salient edge features extraction module. O2OGM: one-to-one guidance module. FF: feature fusion. ISFCREM: improved salient object features compression and reinforcement extraction module.

Methods [12,13,14] pointed out that the shape of CNN pyramid structure usually has a larger space and contains detailed low-order information in the shallow stage, while the deeper stage will contain more high-order information which can accurately highlight the location information of the positioning target. The U-shape structure [15,16] is used in the salient object detection framework [17,18,19] designed by fusing high-order and low-order information, which has caught global low-high level information by uniting different phase features. It constructs the feature fusion path from top to bottom in the classification network which means the low-level semantic information can better fuse the high-level semantic information so as to build a rich feature map.

Recently, based on the U-shape structure designed a new network that is EGNet [20] for salient objection detection which benefit from the rich edge information and location information in salient edge features. The EGNet can help locate salient objects more efficiently. Most of images utilize image blocks as input then using context information and end-to-end deep network structure to output salient map based on each pixel in the image region. After fusing edge information, the local edge information of the object is more prominent and it is used to refine the high-order semantic information. At the same time, the complete consistency between the edge information and the object information is highlighted. However, we found that the original convolution network can be compressed which received the influence of SqueezeNet [21] architecture when fusing the edge information and the salient map information of the context. We changed the convolution design on the premise of nearly keeping the receptive field unchanged

and used the compressed convolution to reduce the parameters of the model. We also find that on the basis of the U-shape structure paying attention to the relationship between the channel can automatically learn the importance of different features according to different channel features so introducing channel attention mechanism is to make up for the loss of the details of the feature map caused by the compression of the model.

We focus the balanced compensation between the model compression of net and the precise of feature maps to save storage space on the premise of almost keeping the accuracy of the original model unchanged. Then we introduce the attention mechanism to focus on the weight attention of the different model channels in order to focus more on salient feature maps and make up for the loss caused by compression. In summary, this paper makes three major contributions:

- We design a new model structure with reduced model size which using smaller CNN architectures are more feasible to deploy on FPGAs and other hardware with limited memory and require less communication across servers during distributed training.

- We introduce the channel attention mechanism to increase the feature focus degree between different channels when the high-order semantic information and low-order semantic information are fused in the up sampling to improve the focus degree of salient map detection offsetting the loss caused by compression.

- We use a new optimization strategy for fast convergence instead of traditional optimization strategy which would possibly lead to local minimums only half of the training time compared with original model.

On the whole, we introduced the related works of the current development method in Sec 2. Next we proposed our network structure and concisely introduced the related modules respectively in Sec 3. Sec 4 is the analysis of results compared with others experiments. Finally we concluded our work in Sec 5.

2. RELATED WORKS

Recently, benefiting from the powerful feature extraction capability of CNNs, most of the traditional salient detection methods based on hand-crafted features [22] have been gradually surpassed.

In recent years, the commonly used salient object detection methods based on deep neural network are mostly optimized from the following three aspects: the method based on boundary enhancement, the method based on semantic enhancement, and the method based on boundary with semantic enhancement. Boundary based enhancement is to obtain more boundary information by enhancing the low-level features of depth features to better locate the significant target boundary. Typical algorithms include ELD [23] algorithm proposed by Tai et al., KSR [24] algorithm proposed by Wang et al., DCL [25] algorithm proposed by Li et al., and DSS algorithm proposed by Hou et al. Semantic enhancement is to obtain rich semantic information from high-level features to better locate the salient target and make the salient target more prominent. Typical algorithms include R-FCN [26] algorithm proposed by Dai et al., CPD [27] algorithm proposed by Wu et al., and PoolNet [28] algorithm proposed by Liu et al. These methods can accurately locate the position of salient object. However, the semantics contained in high-level features are enhanced but sometimes the boundary is blurred or multiple salient object is overlaid. It is easy to cause ambiguity of salient object while only semantic enhancement or boundary enhancement. In order to overcome the shortcomings of the two questions, some

researches enhance the boundary and semantics at the same time which obtain good salient object information and contour information to improve the performance of salient object detection. Typical algorithms include Amulet algorithm proposed by Zhang et al. And BDMPM [29] algorithm proposed by Zhang et al.

Compared with the above network structure design, the popular edge information fusion strategy of EGNet makes up for the salient graph information. Through an end-to-end model design, we can better capture the salient graph information. We explore the model structure

design and use the model compression strategy to reduce the parameters improving the model inference speed while maintaining the same sensing field using smaller CNNs. We introduce the channel attention mechanism to focus on the focusing degree between different map channels to enhance the correlation between channels to better make up for the loss caused by using model compression and use a new optimization strategy to converge the global loss minimum more quickly.

3. NETWORK ARCHITECTURE

The overall architecture is shown in Figure 1. In this section, we begin by describing the compression module in Sec. 3.1, then introduce the channel attention mechanism module in Sec. 3.2, introduce our improved salient object features compression and reinforcement extraction module (ISFCREM) in Sec. 3.3, and finally introduce the new AdaX optimizer in Sec. 3.4.

We also use the VGG [30] network as backbone, followed by the original EGNet we obtain six side features extracting Conv1-2, Conv2-2, Conv3-3, Conv4-3, Conv5-3, Conv6-3. Because the Conv1-2 is too close to the input and the receptive field is too small, we leverage the widely used architecture U-shape from Conv2-2 to Conv6-3 to generate more robust salient object features. But we use fire module instead of original up sampling layers to compress the model. Before adapt a single convolution layer to convert the feature maps to the single-channel prediction mask we use squeeze-and-excitation module [31] (SE module) to study the important weights of different channels which will enhance important and restrain unimportant channel features automatically. Then use the Conv2-2 which preserves better edge information to guide the Convi-3, i∈ [3,6]. Finally, after fuse the O2OGM feature maps to use cross-entropy loss to compute the final sum loss between PSFEM and O2OGM. More details of these modules can be found in original EGNet paper.

3.1. Fire Module

In the model space, we explore the convolution structure and compress the size of the model. The purpose is to achieve the compression of the model parameters while maintaining the same receptive field with smaller convolution kernel. The core idea of SqueezeNet design is introducing the fire module. The design architecture is shown in Figure 2, using 1×1 convolution kernel can limit the number of channels and enhance the abstract expression ability of convolution network which is equivalent to using line sense machine. A larger activation graph is provided for convolution by postponing the down sampling operation which retains more information and can provide higher classification accuracy. The design scale is shown in Figure 3.

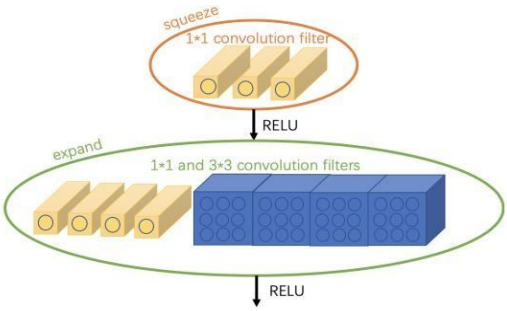

Figure 2. The squeeze and expand Module architecture

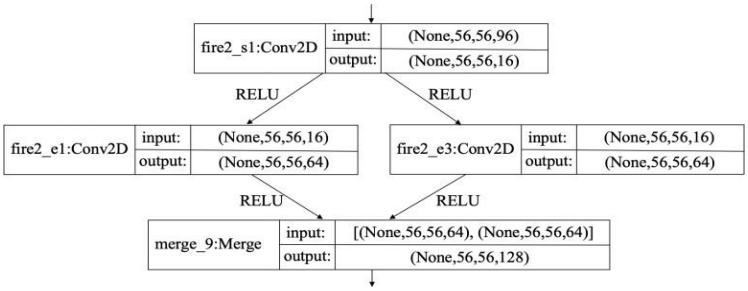

Figure 3. The detailed proportion of squeeze and expand module.

3.2. Squeeze-and-Excitation Module

We explore the channel attention mechanism in order to improve the focus of the feature graph after up sampling on the U-shape structure decoder. We introduce the SE module to allocate the available feature resources to the most valuable information in an input semaphore and enhance the expression ability of the network by aggregating the attention weight ratio between different channels.

3.3. Our Improved Salient Features Compression Reinforcement Extraction Module

Our improved module (ISFCREM) use 1×1 convolution kernel on the squeeze portion to reduce dimension and combining 3×3 with 1 ×1 convolution kernel of expand portion keep the ratio of 1:1 to increase dimension which not only decreases the kernel parameters caused by the 3 × 3 convolution kernel but also retain the same received sensation. The number of input channels in the fire module keeps the same receptive field as the original convolution to the greatest extent and reduces the count of parameters.

Secondly, concatenate the SE module. SE module first compresses the global spatial information by using the global average pooling to generate the statistical information between channels for a channel descriptor so that the following network layer can obtain the global receptive field information. In the second step, one fully connected layer is used to compress the number of channels to reduce the amount of computation and then the second fully connected layer is used to recover the number of channels. The two fully connected layers leverage the different channel correlation to enhance important features and weaken unnecessary features. Finally, the sigmoid activation function is used to enlarge and shrink the range and the scale factor is fused into the original feature map as the weight ratio then add directly for enhancing the weight of important feature map.

Our improved module (ISFCREM) compress the model while maintaining maximum accuracy.

3.4. The Optimizer of AdaX

Compared with the Adam optimizer [32], we use a new optimizer that is AdaX [33] optimizer. The main formulation of updating parameters can be seen as Table 1.

Table 1. The different formulation of optimization strategy between Adam and AdaX optimizer

$g_t = \nabla_\theta L(\theta_t)$ (1) $m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t$ (2) $v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2$ (3) $\hat{m}_t = m_t / (1-\beta_1^t)$ (4) $\hat{v} = v_t / (1-\beta_2^t)$ (5) $\theta_t = \theta_{t-1} - \alpha_t \hat{m} / \sqrt{\hat{v}} + \epsilon$ (6)	$g_t = \nabla_\theta L(\theta_t)$ (7) $m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t$ (8) $v_t = (1+\beta_2) v_{t-1} + \beta_2 g_t^2$ (9) $\hat{v} = v_t / ((1+\beta_2)^t - 1)$ (10) $\theta_t = \theta_{t-1} - \alpha_t m_t / \sqrt{\hat{v}} + \epsilon$ (11)
(a）The mainly updating formula (1-6) of AdaXoptimizer	（b）The mainly updating formula (7-11) of Adam optimizer

The difference of the two optimizer is that AdaX optimizer remove the offset correction of momentum (4) firstly. Secondly, AdaX optimizer can adapt its learning rate by itself with exponential longterm memory. As time goes on, the decay strategy of Adam optimizer learning rate may not converge a constant but we hope that the learning rate to be a constant because the gradient becomes smaller and the training itself tends to be stable in the later stage of training. So the significance of correcting the learning rate is not great and the correction strength of the leaning rate should become smaller that should be better a constant. As the formulation can be seen that AdaX optimizer can better convergence a constant that make the loss convergence smoother than Adam optimizer. Our training experiment results using three different optimizer for a comparison can be shown as Figure 4.

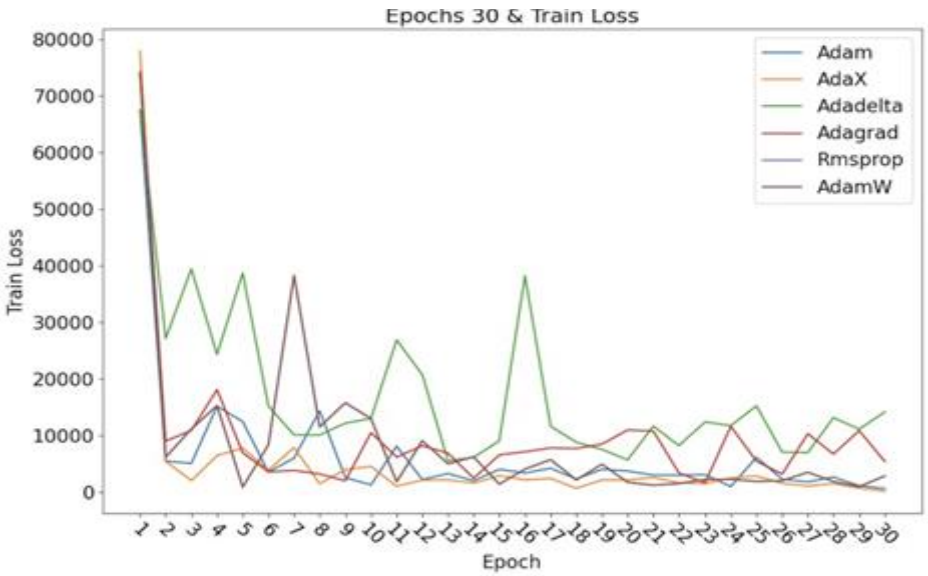

Figure 4. The comparison between Adadelta, Adam, Adagrad, Rmsprop, AdamW and AdaX optimizer based on Resnet-50 as backbone after training 30 epochs on DUTS-TE salient object detection dataset.

4. EXPERIMENTS

4.1. Implementation Details

We train the improved model on DUTS-TS [34] dataset in our device which contains two P40 GPUs having 24GB video random access memory each other. We also use VGG-16 and Resnet-50 [35] as backbone for a comparison respectively. Our code is implied by PyTorch. We set the weights of newly added convolution layers are initialized randomly with a truncated normal (σ=0.01), and the biases are initialized to 0. We use AdaX as the optimizer. The hyperparameters are set as followed: learning rate=5e-5, weight decay=0.0005, momentum = 0.9, loss weight for each side output is equal to 1. A back propagation is processing for each of the ten images. We do not use the validation dataset during training. We train our model 18 epochs and divide the learning rate by 10 after 9 epochs. We apply the official model to train original experiment as comparison. The results can be shown as Figure 5.

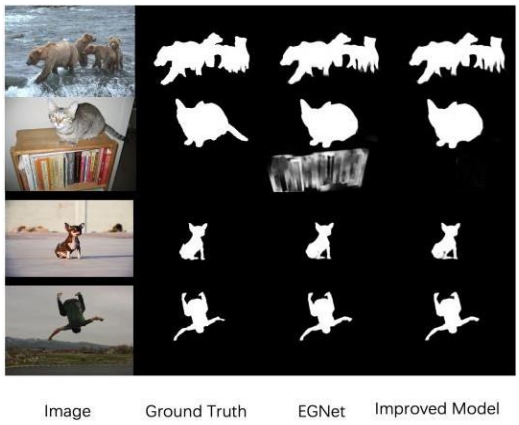

Figure 5. The result salient map of improved model based on the Resnet-50 can be seen as a contrast with the original net. As we can see, the salient object can be focus more of its construction when we imported the SE module in the second line feature maps.

4.2. Datasets and Evaluation Criteria

We test the improved model on six widely public benchmark datasets: DUTOMRON [36], DUTS-TE, ECSSD, PASCAL-S [37], HKU-IS [38] and SOD [39] that altogether nearly contains 30000 meaningful semantic images with various complex scenes. DUTS dataset contains 10553 original images, saliency images and corresponding edge images. DUTOMRON dataset contains 5168 original images and salient feature images respectively. Images in this dataset contain one or more salient objects with a relatively complex background. SOD, ECSSD, PASCAL-S and HKU-IS dataset includes 300, 1000, 850, 4447 original images and salient feature images respectively. DUTS-TE dataset includes 5017 original images and salient feature images respectively. The dataset of DUTS is the largest salient object detection stand which include more challenged images with various locations and scales.

We use four metrics, F-Value, mean absolute error (MAE), S-measure [40] to compare with baseline model. F-value is a mean of average precision and average recall, the formula can be described as Equation (12):

$$F_\beta = \frac{(1+\beta^2)*\text{Precision}*\text{Recall}}{\beta^2*\text{Precision}+\text{Recall}} \tag{12}$$

where β^2 is set to 0.3 as done in previous work to weight precision more than recall. The MAE score indicates how similar a saliency map S is compared to the ground truth G and the formula as Equation (13):

$$MAE = \frac{1}{W*H}\sum_{x=1}^{W}\sum_{y=1}^{H}|S(x,y) - G(x,y)| \quad (13)$$

where W and H denote the width and height of S, respectively. The IOU can be described as the intersection-over-union between the predicted map S and ground truth G and the formula as Equation (14):

$$IOU = \frac{S \cap G}{S \cup G} \quad (14)$$

S-measure focuses on evaluating the structural information of salient maps. It is closer to the human visual system than F-measure. We include S-measure for a more comprehensive evaluation for comparison. S-measure could be computed as Equation (15):

$$S = \gamma S_o + (1-\gamma)S_r \quad (15)$$

4.3. Evaluation Results

First of all, we compare the traditional mainstream method with our improved model on ESSCD, PASCAL-S, DUTOMRON, HKU-IS, SOD and DUTS-TE databases respectively, and the comparison results are shown in Table 2. We draw the PR curve on DUTOMRON, PASCAL-S, and DUTS-TE datasets which are shown on Figure 6. Next, we also compared the results of IOU, Inference Speed between EGNet and our improved model. The result shown as Table 3 and Table 4 respectively.

Table 2. Quantitative comparison including max F-measure, MAE, and S-measure over six widely used datasets. "-" denotes that corresponding methods are trained on that dataset. "*" means methods using pre-processing or post-processing. Compared with traditional methods that our improved model evaluation result is nearly the same as EGNet but more better than others results. The best results are marked in red and the second best results are marked blue respectively.

	ECSSD			PASCAL-S			DUTOMRON			HKU-IS			SOD			DUTS-TE		
	MaxF	MAE	S	MaxF	MAE	S	MaxF	MAE	S	MaxF	MAE	S	MaxF	MAE	S	MaxF	MAE	S
									VGG-based									
DCL*	0.898	0.082	0.863	0.806	0.114	0.790	0.735	0.094	0.742	0.895	0.061	0.860	0.831	0.133	0.749	0.784	0.080	0.785
DSS*	0.905	0.066	0.881	0.822	0.101	0.795	0.760	0.073	0.766	0.901	0.052	0.879	0.835	0.125	0.745	0.812	0.066	0.814
MSR	0.904	0.059	0.875	0.840	0.082	0.801	0.791	0.072	0.767	0.909	0.045	0.853	0.840	0.109	0.757	0.825	0.064	0.810
NLDF	0.905	0.066	0.874	0.821	0.098	0.805	0.752	0.080	0.751	0.900	0.049	0.878	0.837	0.125	0.755	0.814	0.067	0.807
RAS [41]	0.917	0.060	0.885	0.832	0.104	0.799	0.784	0.062	0.793	0.911	0.048	0.886	0.845	0.132	0.762	0.802	0.060	0.826
ELD*	0.866	0.084	0.840	0.772	0.123	0.756	0.736	0.093	0.744	0.845	0.073	0.825	0.761	0.134	0.706	0.745	0.093	0.750
DHS [42]	0.904	0.064	0.885	0.826	0.090	0.807	-	-	-	0.891	0.052	0.870	0.825	0.129	0.749	0.814	0.064	0.809
RFCN*[43]	0.897	0.096	0.854	0.825	0.120	0.799	0.747	0.096	0.753	0.894	0.080	0.860	0.807	0.163	0.731	0.784	0.088	0.782
UCF [44]	0.906	0.082	0.886	0.820	0.126	0.804	0.732	0.133	0.749	0.886	0.075	0.872	0.799	0.164	0.760	0.770	0.114	0.778
Amulet	0.913	0.063	0.895	0.827	0.094	0.820	0.735	0.085	0.771	0.887	0.050	0.888	0.797	0.148	0.753	0.773	0.076	0.798
C2S [45]	0.910	0.057	0.892	0.843	0.082	0.839	0.760	0.074	0.785	0.899	0.045	0.887	0.823	0.124	0.765	0.811	0.060	0.821
PAGR	0.924	0.062	0.888	0.849	0.088	0.817	0.772	0.072	0.752	0.917	0.047	0.887	0.842	0.146	0.718	0.853	0.055	0.824
EGNet	0.942	0.045	0.914	0.863	0.075	0.846	0.824	0.054	0.813	0.927	0.036	0.912	0.870	0.111	0.789	0.880	0.045	0.868
Ours	0.936	0.045	0.909	0.855	0.081	0.837	0.815	0.060	0.815	0.925	0.037	0.903	0.854	0.114	0.778	0.869	0.046	0.865
									ResNet-based									
SRM*[49]	0.914	0.056	0.897	0.836	0.085	0.832	0.770	0.071	0.775	0.908	0.044	0.888	0.842	0.128	0.743	0.824	0.059	0.824
DGRL[52]	0.922	0.044	0.908	0.844	0.074	0.840	0.775	0.062	0.793	0.912	0.038	0.898	0.845	0.105	0.772	0.828	0.050	0.837
PiCANet*[33]	0.934	0.046	0.914	0.865	0.078	0.851	0.822	0.062	0.809	0.922	0.045	0.906	0.861	0.101	0.788	0.862	0.051	0.850
EGNet	0.944	0.042	0.917	0.870	0.074	0.853	0.843	0.053	0.817	0.936	0.031	0.919	0.892	0.097	0.806	0.892	0.040	0.876
Ours	0.938	0.043	0.913	0.858	0.074	0.840	0.818	0.058	0.815	0.928	0.035	0.905	0.865	0.116	0.774	0.858	0.044	0.867

Table 3. The evaluation of IOU between different datasets compared EGNet with our improved model.

	ECSSD	PASCAL-S	DUTOMRON	HKU-IS	SOD	DUTS-TE
VGG-based						
EGNet	0.867	0.750	0.697	0.843	0.654	0.767
Ours	0.846	0.728	0.658	0.828	0.639	0.737
Resnet-based						
EGNet	0.878	0.755	0.701	0.858	0.692	0.782
Ours	0.854	0.735	0.654	0.832	0.637	0.740

Table 4. The evaluation of inference speed between different datasets compared EGNet with our improved model. The inference speed of our improved model is faster than EGNet.

	PASCAL-S (seconds/sheet)	DUTOMRON (seconds/sheet)	HKU-IS (seconds/sheet)	SOD (seconds/sheet)	DUTS-TE (seconds/sheet)
VGG-based					
EGNet	0.476	0.126	0.244	0.117	0.113
Ours	0.448	0.142	0.223	0.090	0.122
Resnet-based					
EGNet	0.128	0.131	0.270	0.180	0.129
Ours	0.113	0.076	0.198	0.103	0.078

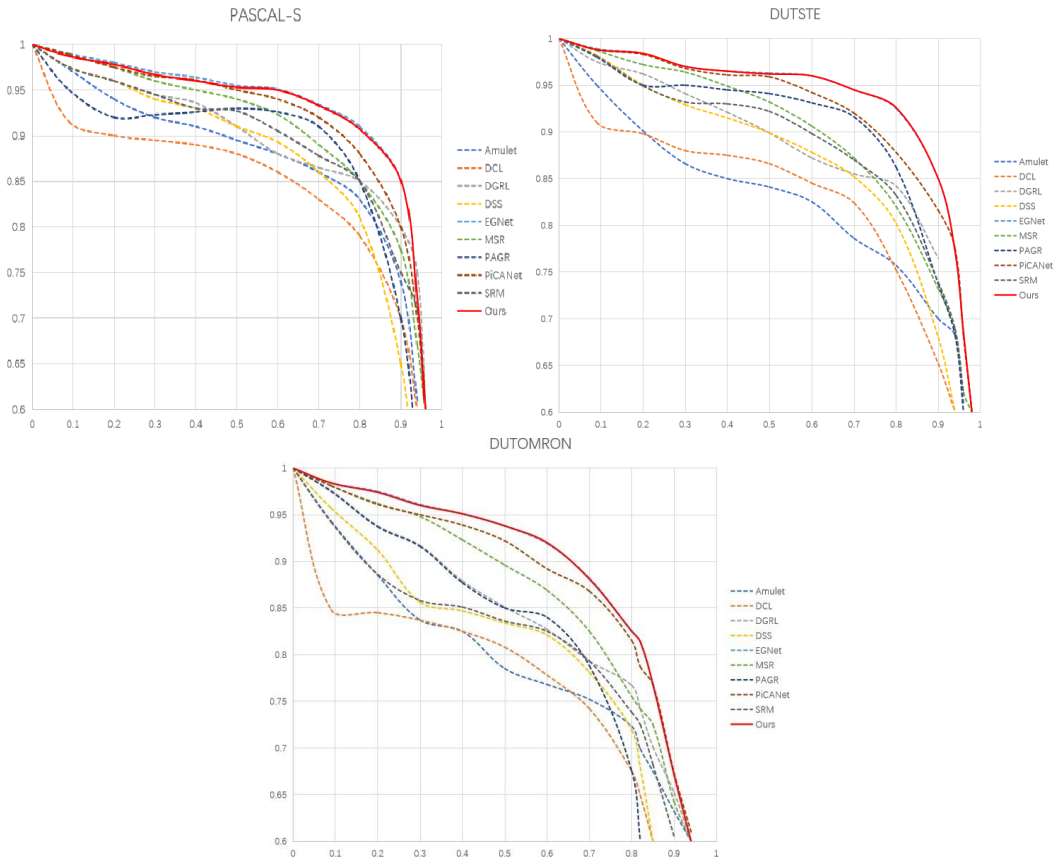

Figure 6. The Precision and Recall for PASCAL-S, DUTSTE, DUTOMRON datasets respectively. Our improved model is better than others and nearly similar as the state-of-the-art EGNet.

5. CONCLUSIONS

In this paper, we use smaller CNNs compressed the model size based on the EGNet structure firstly. The parameters are reduced from 108068426 to 30175602 (111692618 to 33799794) and the model size from 412MB to 115MB (426MB to 129MB) which computed on linux system using VGG-16 (Resnet-50) as the backbone respectively. It is a third of the original size. But we found that the accuracy is lower 3.0+ compared with the-state-of-art that is EGNet. For a remedy, we introduce channel attention to minimize gaps. The results shown that our improved model is close to the-state-of-the-art. We also explore a new optimizer for better and faster convergence a global minimum loss. The experiment shown that its test result is crowded on the Adam optimizer which is the best optimizer widely used but nearly half of training time.

In the future work, we want to explore a larger proportion of compression models by means of model weight pruning and knowledge distillation. In terms of model accuracy, we will continue to explore the edge structure extraction of refined feature map to make the segmented region more effective. At the same time, we will explore new self-attention structure mechanisms, such as introducing the latest transformer mechanism to capture global context information and improve segmentation accuracy.

ACKNOWLEDGEMENTS

This work is supported by CERNET Innovation Project (NGII20190625) and The Inner Mongolia Natural Science Foundation of China under Grant No.2021MS06016.

REFERENCES

[1] L SeunghoonHong, TackgeunYou, SuhaKwak, and Bohyung Han. Online tracking by learning discriminative saliency map with convolutional neural network. In ICML, pages 597–606, 2015. 1

[2] Ming-Ming Cheng, Fang-Lue Zhang, Niloy J Mitra, Xiaolei Huang, and Shi-Min Hu. Repfinder: finding approximately repeated scene elements for image editing. ACM TOG, 29(4):83, 2010. 1

[3] Celine Craye, David Filliat, and Jean-Francois Goudou. Environment exploration for objectbased visual saliency learning. In ICRA, pages 2303–2309, 2016. 1

[4] Laurent Itti, Christof Koch, and Ernst Niebur. A model of saliencybased visual attention for rapid scene analysis. IEEE TPAMI, 20(11):1254–1259, 1998. 1

[5] Tie Liu, Zejian Yuan, Jian Sun, Jingdong Wang, Nanning Zheng, Xiaoou Tang, and Heung-Yeung Shum. Learning to detect a salient object. IEEE TPAMI, 33(2):353–367, 2011. 1, 5, 6, 7

[6] Dominik A Klein and Simone Frintrop. Center-surround divergence of feature statistics for salient object detection. In ICCV, 2011. 1

[7] Federico Perazzi, Philipp Kra henbu hl, Yael Pritch, and Alexander Hornung. Saliency filters: Contrast based filtering for salient region detection. In CVPR, pages 733–740, 2012. 1, 2

[8] Ali Borji and Laurent Itti. Exploiting local and global patch rarities for saliency detection. In CVPR, pages 478–485, 2012. 1

[9] Huaizu Jiang, Jingdong Wang, Zejian Yuan, Yang Wu, Nanning Zheng, and Shipeng Li. Salient object detection: A discriminative regional feature integration approach. In CVPR, pages 2083–2090, 2013. 1, 2

[10] Qiong Yan, Li Xu, Jianping Shi, and Jiaya Jia. Hierarchical saliency detection. In CVPR, pages 1155–1162, 2013. 1, 5, 8

[11] Ming Cheng, Niloy J Mitra, Xumin Huang, Philip HS Torr, and Song Hu. Global contrast based salient region detection. IEEE TPAMI, 2015. 1, 2, 8

[12] Qibin Hou, Ming-Ming Cheng, Xiaowei Hu, Ali Borji, Zhuowen Tu, and Philip Torr. Deeply supervised salient object detection with short connections. IEEE TPAMI, 41(4):815–828, 2019. 1, 2, 3, 5, 6, 7, 8

[13] Zhiming Luo, Akshaya Kumar Mishra, Andrew Achkar, Justin A Eichel, Shaozi Li, and Pierre-Marc Jodoin. Nonlocal deep features for salient object detection. In CVPR, 2017. 1, 2, 7, 8

[14] Pingping Zhang, Dong Wang, Huchuan Lu, Hongyu Wang, and Xiang Ruan. Amulet: Aggregating multilevel convolutional features for salient object detection. In ICCV, 2017. 1, 2, 6, 7, 8

[15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241, 2015. 1

[16] Tsung-Yi Lin, Piotr Dolla r, Ross B Girshick, Kaiming He, Bharath Hariharan, and Serge J Belongie. Feature pyramid networks for object detection. In CVPR, 2017. 1, 2, 3, 4, 6

[17] Guanbin Li, Yuan Xie, Liang Lin, and Yizhou Yu. Instance-level salient object segmentation. In CVPR, 2017. 1, 5, 6, 7, 8

[18] TiantianWang, LiheZhang, ShuoWang, HuchuanLu, Gang Yang, Xiang Ruan, and Ali Borji. Detect globally, refine locally: A novel approach to saliency detection. In CVPR, pages 3127–3135, 2018. 1, 3, 6, 7, 8

[19] ibin Hou, Jiang-Jiang Liu, Ming-Ming Cheng, Ali Borji, and Philip HS Torr. Three birds one stone: A unified framework for salient object segmentation, edge detection and skeleton extraction. arXiv preprint arXiv:1803.09860, 2018. 1, 7, 8

[20] Zhao J, JJ Liu, Fan D P, et al. EGNet: Edge Guidance Network for Salient Object Detection[C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, 2020.

[21] Iandola F N, Han S, Moskewicz M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size[J]. 2016.

[22] Xiaohui Li, Huchuan Lu, Lihe Zhang, Xiang Ruan, and Ming Hsuan Yang. Saliency detection via dense and sparse reconstruction. In ICCV, pages 2976–2983, 2013. 2

[23] LEE G, TAI Y W, KIM J, et al. Deep saliency with encoded low level distance map and high level features[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, Jun 27-30, 2016. Washington: IEEE Computer Society, 2016: 660-668.

[24] WANG T T, ZHANG L H, LU H C, et al. Kernelized subspace ranking for saliency detection[C]//LNCS 9912: Proceedings of the 14th European Conference on Computer Vision, Amsterdam, Oct 11- 14, 2016. Berlin, Heidelberg: Springer, 2016: 450-466.

[25] LI G B, YU Y Z. Deep contrast learning for salient object detection[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, Jun 27- 30, 2016. Washington: IEEE Computer Society, 2016: 478-487.

[26] DAI J F, LI Y, HE K, et al. R-FCN: object detection via region- based fully convolutional networks[C]//Proceedings of the 30th International Conference on Neural Information Processing Systems, Barcelona, Dec 5, 2016. Massachusetts: MIT Press, 2016: 379-387.

[27] WU Z, SU L, HUANG Q M, et al. Cascaded partial decoder for fast and accurate salient object detection[C]//Proceedings of the 2019 IEEE Conference on Computer Vision and Pat- tern Recognition, Long Beach, Jun 16-20, 2019. Washington: IEEE Computer Society, 2019: 3907-3916.

[28] LIU J J, HOU Q B, CHENG M M, et al. A simple pooling-based design for real-time salient object detection[C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, Jun 16-20, 2019. Piscataway: IEEE, 2019: 3917-3926.

[29] ZHANG L, DAI J, LU H C, et al. A bi-directional message passing model for salient object detection[C]//Proceedings of the 2018 IEEE Conference on Computer Vision and Pat- tern Recognition, Salt Lake City, Jun 18-22, 2018. Washing- ton: IEEE Computer Society, 2018: 1741-1750.

[30] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.

[31] Jie H, Li S, Gang S. Squeeze-and-Excitation Networks[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2018.

[32] Kingma D, Ba J. Adam: A Method for Stochastic Optimization[J]. Computer Science, 2014.

[33] Li W, Zhang Z, Wang X, et al. AdaX: Adaptive Gradient Descent with Exponential Long Term Memory[J]. 2020.

[34] Xiaoning Zhang, Tiantian Wang, Jinqing Qi, Huchuan Lu, and Gang Wang. Progressive attention guided recurrent network for salient object detection. In CVPR, pages 714–722, 2018. 1, 3, 6, 7, 8

[35] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016.

[36] Chuan Yang, Lihe Zhang, Huchuan Lu, Xiang Ruan, and Ming-Hsuan Yang. Saliency detection via graph-based manifold ranking. In CVPR, pages 3166–3173, 2013.

[37] Yin Li, Xiaodi Hou, Christof Koch, James M Rehg, and Alan L Yuille. The secrets of salient object segmentation. In CVPR, pages 280–287, 2014.

[38] Guanbin Li and Yizhou Yu. Visual saliency based on multi- scale deep features. In CVPR, pages 5455–5463, 2015.

[39] David Martin, Charless Fowlkes, Doron Tal, and Jitendra Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In ICCV, volume 2, pages 416–423, 2001.

[40] Cheng M M, Fan D P. Structure-Measure: A New Way to Evaluate Foreground Maps[J]. International Journal of Computer Vision, 2021, 129(9):2622-2638.

[41] Shuhan Chen, Xiuli Tan, Ben Wang, and Xuelong Hu. Reverse attention for salient object detection. In ECCV, pages 234–250, 2018.

[42] Nian Liu and Junwei Han. Dhsnet: Deep hierarchical saliency network for salient object detection. In CVPR, pages 678–686, 2016.
[43] Linzhao Wang, Lijun Wang, Huchuan Lu, Pingping Zhang, and Xiang Ruan. Saliency detection with recurrent fully convolutional networks. In ECCV, 2016.
[44] Pingping Zhang, Dong Wang, Huchuan Lu, Hongyu Wang, and Baocai Yin. Learning uncertain convolutional features for accurate saliency detection. In ICCV, pages 212–221. IEEE, 2017.
[45] Xin Li, Fan Yang, Hong Cheng, Wei Liu, and Dinggang Shen. Contour knowledge transfer for salient object detection. In ECCV, pages 355–370, 2018.

AUTHORS

Ming Ma is a lecturer of Inner Mongolia University, graduated from Pai Chai University in Korea with a Ph. in 2009. He joined the School of Computer Science of Inner Mongolia University as a teacher from 2013.

Bin Zhang is currently a master of computer technology, Inner Mongolia University, under the super-vision of Lecturer Ming Ma. His research interests include deep learning, image processing, and computer vision.

Yang Wu is a doctoral student at the College of Computer Science of Inner Mongolia University. He is mainly interested in action recognition and video classification tasks.

Xiaojing Zhang is currently a graduate student with College of Electronic Information Technology, Inner Mongolia University, under the super-vision of Lecturer Ming Ma. Her research interests include deep learning, image processing, and computer vision.

DEMPSTER-SHAFER AND MULTI-FOCUS IMAGE FUSION USING LOCAL DISTANCE

Ias Sri Wahyuni[1] and Rachid Sabre[2]

[1]University of Burgundy, Gunadarma University
[2]Laboratory Biogéosciences CNRS,
University of Burgundy/Agrosup Dijon, France

ABSTRACT

In this article, we give a new method of multi-focus fusion images based on Dempster-Shafer theory using local variability (DST-LV). Indeed, the method takes into account the variability of observations of neighbouring pixels at the point studied. At each pixel, the method exploits the quadratic distance between the value of the pixel I (x, y) of the point studied and the value of all pixels which belong to its neighbourhood. Local variability is used to determine the mass function. In this work, two classes of Dempster-Shafer theory are considered: the fuzzy part and the focused part. We show that our method gives the significant and better result by comparing it to other methods.

KEYWORDS

Multi-focus-images, Dempster-Shafer Theory, local distance, fusion images.

1. INTRODUCTION

Image fusion involves combining images from the same scene to produce a single image containing more information and detail found on at least one of the input images. Thus, image fusion can reduce uncertainty and minimize redundancy on the output image as well as maximize particular relevant information. This paper deals with the fusion of multi focus images caused by the limited depth of field of optical lenses in cameras. Thus, it is not possible to obtain an image containing all the relevant objects in a scene but with different focusing, one can obtain several images where each one contains a clear object and the others blurry. The image fusion method is used to get all objects in focus on a single image.

In the literature, we find several methods of fusion of multifocal images. These methods can be two types, spatial methods and multiscale methods. Those of the spatial domain directly concern the pixels of the source images as well as their neighbourhoods. Moreover, the fusion methods such as the mean, the principal component analysis (PCA) [1], the maximum selection rule, the methods based on the two-sided gradient [2] and the method based on the filter and Guided images (GIF) [3] and the maximum selection rule are considered as spatial approaches. The flaw of spatial domain approaches is that they cause spatial distortion in the fused image. On the other hand, the fusion by the methods at several scales is carried out on the source images after having decomposed them into several scales. As examples of these methods we cite among others: discrete wavelet transform (DWT) [4] - [7], the fusion of Laplacian pyramidal images [8] - [14], the discrete cosine transform with calculation of the variance (DCT + var) [15], the method based on the detection of salience (SD) [16].

David C. Wyld et al. (Eds): CSTY, SIGI, MaVaS - 2021
pp. 59-73, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.112206

The paper [17] show that, from the evidence point of view where the fusion decreases the imprecision and the uncertainty by using the redundancy and the complementary information of the source image. This means that evidence of weak inputs gives the best estimate. Shafer first proposed the evidence theory in the 1970s, based on Dempster's research. The advantage of the Dempster-Shafer theory (DST) is without a priori and without preference, because of the unavailability of information, which implies an indeterminacy, as detailed in [18] and [19]. This theory is very successful in many applications, including image segmentation [20], [21], pattern classification [22], [23], object recognition [24], medical imaging. [25], fusion of sensors [26].

This paper proposes the fusion of multi-focus images using Dempste-Shafer theory based on the following information: the variability between each pixel and its neighbours. This variability is calculated from the quadratic distance between the value of the pixel I (x, y) and the value of all neighbouring pixels. It is called "local variability". Local variability as a measure can detect the sharp intensity of the image such as the edge. This method also takes into account the behaviour of neighbouring pixels and preserves the edge.

The organization of this paper is as follows: Section 2 gives details of Dempster-Shafer evidence theory. Section 3 gives the definition of local variability and its use. Section 4 presents the proposed method in detail. Section 5 provides the definition of the evaluation metrics used in this article. The experimental study using different images and the comparison between the proposed method and other methods are provided in section 6. The section 7 is reserved for the conclusions and perspective of this work.

2. DEMPSTER-SHAFER EVIDENCE THEORY

Define Θ as the set of hypotheses for a problem domain, called frame of discernment. Let *m* the function defined from 2^{Θ} to $[0,1]$ where 2^{Θ} be the set of all subsets Θ :

$$2^{\Theta} = \{A | A \subseteq \Theta\}. \tag{1}$$

The function *m* is called a basic probability assignment whenever

$$\mathrm{m}(\varnothing) = 0 \text{and} \sum_{A \subset \Theta} \mathrm{m}(A) = 1. \tag{2}$$

Where $m(A)$ is the measure of belief that is attributed to A. According to [27], $m(A)$ is the degree of evidence supporting the claim that a specific element of Θ belongs to the set *A*, but not to any special subset of complementary of A. If A is a element of Θ such that $m(A) > 0$ then *A* is called the focal element of m. The belief measure is given by using *m* as follows: $Bel: 2^{\Theta} \mapsto [0,1]$:

$$Bel(A) = \sum_{B \subset A} m(B). \tag{3}$$

The paper [28] define the plausibility measure $Pl: 2^{\Theta} \mapsto [0,1]$:

$$Pl(A) = \sum_{A \cap B \neq \varnothing} m(B) = 1 - Bel(\overline{A}). \tag{4}$$

$Bel(A)$ measures the degree of evidence that the element in question belongs to the set A as well as to the various special subsets of A. An important aspect of Dempster Shafer Theory (DST) is the aggregation of evidence given by different sources, see [17]. If two mass function m_1 and m_2

applied at distinct items of evidence are such that $m_1(B) > 0$ and $m_1(C) > 0$ for some non-disjoint subsets B and C of Θ, then they are combinable by means of Dempster's rule. [29], [30]. The combination (joint mass) of two sets of masses m_1 and m_2 is defined as follows

$$m_1 \oplus m_2(\varnothing) = 0 \tag{5}$$

$$m_1 \oplus m_2(A) = \frac{\sum_{B\cap C=A} m_1(B)m_2(C)}{1-\sum_{B\cap C=\varnothing} m_1(B)m_2(C)} \tag{6}$$

Equation (6) becomes

$$m_1 \oplus m_2(A) = \frac{\sum_{B\cap C=A} m_1(B)m_2(C)}{\sum_{B\cap C\neq\varnothing} m_1(B)m_2(C)}. \tag{7}$$

If the mass on a subset A is zero does not mean that the set is impossible, simply that we are not capable of assigning a level precisely to A (see [31]), since we could have non-zero masses on subsets of A, which would lead us to $Bel(A) \neq 0$.

3. LOCAL DISTANCE

This paper considerate the information of the neighbouring pixels of the fixed pixel. Indeed, at each pixel $I(x,y)$, the method uses the quadratic difference called local variability between the value of this pixel $I(x,y)$ and the value of its neighbours. The idea comes from the fact that the variability of the value in blurred region is smaller than the variability of the value in the focused region, the proof of this assertion is given in [32]. The neighbour of a pixel (x,y) used in this paper, with the size`"a" is:

$$(\mathrm{x}+\mathrm{i},\mathrm{y}+\mathrm{j})\mathrm{where}\mathrm{i} = -\mathrm{a},-\mathrm{a}+1,\cdots,\mathrm{a}-1,\mathrm{a}\mathrm{and}\mathrm{i} = -\mathrm{a},-\mathrm{a}+1,\cdots,\mathrm{a}-1,\mathrm{a}$$

For example the neighbor with the small size (`"a" $= 1$) contains: $(x-1,y-1)$, $(x-1,y)$, $(x-1,y+1)$, $(x,y-1)$, $(x,y+1)$, $(x+1,y-1)$, $(x+1,y)$, $(x+1,y+1)$ as we can see in Fig. 1.

	(*x*-1, *y*-1)	(*x*-1, *y*)	(*x*-1, *y*+1)		
	(*x*, *y*-1)	(*x*, *y*)	(*x*, *y*+1)		
	(*x*+1, *y*-1)	(*x*+1, *y*)	(*x*+1, *y*+1)		

Figure 1. Pixel at (x, y) within its neighborhood, a = 1

Cosider p source images $(I_1, I_2, \cdots, I_p)$ with same size $(R \times C)$. Local variability of every source image at pixel (x,y):

$$v_{a,k}(x,y) = \sqrt{\frac{1}{T}\sum_{m=-a}^{a}\sum_{n=-a}^{a}|I_k(x,y) - I_k(x+m,y+n)|^2} \quad (8)$$

where k is the index of k^{th} source image $(k = 1,2,\cdots,p)$.

$$I_k(x+m,y+n) = \begin{cases} I_k(x+m,y+n), & \text{if } 1 \le x+m \le R \text{ and } 1 \le y+n \le C, \\ I_k(x,y), & \text{otherwise} \end{cases}$$

$$T = (2a+1)^2 - card(S)$$

$$S = \{(m,n) \in ([-a,a]^2 - \{0,0\}) \; such\; that\; I_k(x+m,y+n) = I_k(x,y)\}$$

The following proposition that the local variability is small enough where the location is on the blurred area (B_1 *or* B_2). Indeed, we consider, without loss the generality, that we have a focus pixel (x,y) in image I_1 and blurred in image I_2, $((x,y) \in B_2)$

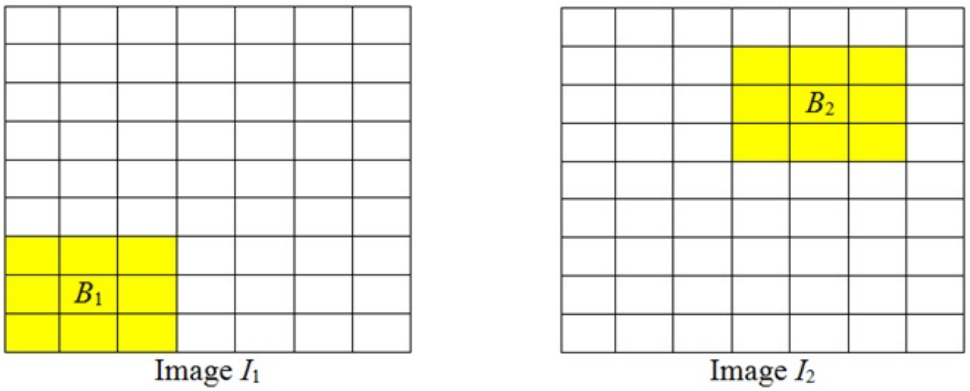

Figure 2. Two multi focus images, the yellow part is blurred area and the white part is clear(focused) area.

The local variability of image I_1 and image I_2 are respectively: $\sqrt{\frac{1}{T}r_1(x,y)}$ and $\sqrt{\frac{1}{T}r_2(x,y)}$, where $r_1(x,y)$ and $r_2(x,y)$ can be written as follow:

$$r_1(x,y) = \sum_{m=0}^{2a}\sum_{n=0}^{2a}|I_1(x,y) - I_1(x+(m-a),y+(n-a))|^2 \quad (9)$$
$$r_2(x,y) = \sum_{m=0}^{2a}\sum_{n=0}^{2a}|I_2(x,y) - I_2(x+(m-a),y+(n-a))|^2 \quad (10)$$

Proposition

Let (x,y) a pixel belongs to blurred area of the image I_2 ($(x,y) \in B_2$), then the local variability on (x,y) in image I_2, is smaller that the local variability on (x,y) in image I_1, ($r_2(x,y) < r_1(x,y)$).

The proof of this proposistion is given in [32].

4. THE PROPOSED METHOD

The difficulty of merging images using Dempster-Shafer theory is to construct the evidential representation of images. In this paper, one information used as the evidential representation images is local variability and two classes are used in the Dempster-Shafer theory. Either a pixel belongs to blurred part ω or it belongs to the focus part $\overline{\omega}$. There is also uncertainty Θ inherent in the theory of evidence. All of this forms the framework for discernment in Θ :

$$\Theta = \{\omega, \overline{\omega}, \theta\} \tag{12}$$

For each pixel one value of evidence for information will be obtained, m.

$$\{m(\omega), m(\overline{\omega}), m(\theta)\} \tag{13}$$

with the condition $m(\omega) + m(\overline{w}) + m(\theta) = 1$.

Suppose there are p original images, $I_1, I_2, \cdots, I_p$, where each image has size $(R \times C)$ with different focus to be fused. The fusion in this work follows 3 steps:

Step 1:

1. To calculate mass function:

For each image where we use different values of size of neighbourhood, $a \in \{1,2,\cdots,10\}$, we define:

$$d'_{a,k}(x,y) = 1 - \frac{v_{a,k}(x,y) - \min\limits_{(x',y')} \left(v_{a,k}(x',y')\right)}{\max\limits_{(x',y')} \left(v_{a,k}(x',y')\right) - \min\limits_{(x',y')} \left(v_{a,k}(x',y')\right)} \tag{14}$$

where k is the k^{th} source image, $k \in \{1,2,\cdots,p\}$ and a is size of neighbourhood of local variability. We set the standard deviation of $d'_{a,k}(x,y) = \sigma_{a,k}(x,y)$,

for (x,y) belongs to ω, we calculate:

$$m_{a,k}(\omega) = \left(1 - \sigma_{a,k}(x,y)\right) d'_{a,k}(x,y) \tag{15}$$

for (x,y) belongs to θ, we calculate:

$$m_{a,k}(\theta) = \sigma_{a,k}(x,y) \tag{16}$$

for (x,y) belongs to $\overline{\omega}$, we calculate:

$$\begin{aligned} m_{a,k}(\overline{\omega}) &= 1 - \left(1 - d'_{a,k}(x,y)\right) \sigma_{a,k}(x,y) - \sigma_{a,k}(x,y) \\ &= \left(1 - d'_{a,k}(x,y)\right)\left(1 - \sigma_{a,k}(x,y)\right) \end{aligned} \tag{17}$$

This method obtains the information whether or not a pixel belongs to the focus area, for this we

use the plausibility of ω which is the sum of the masses of the evidence for ω and the uncertainty θ:

$$Pl_{a,k}(\omega) = m_{a,k}(\omega) + m_{a,k}(\theta)$$

For fusion image at the pixel (x, y), due to ω is a set of pixel on blurred area, we take pixel (x, y) from image k_0 that assigned to minimum $Pl_k(\omega)$, $k = 1,2,\cdots,p$.

Step 2.

For (x, y), we take F_a as fused image with size of neighborhood = a

$$\begin{aligned} F_a(x,y) = I_{k_0}(x,y), \text{where } k_0 \in \{1,2,\cdots,p\} \text{ and } Pl_{a,k_0}(\omega)(x,y) \\ = \min_{k\in\{1,2,\cdots,p\}} \left(Pl_{a,k}(\omega)(x,y)\right). \end{aligned}$$

Step 3.

The proposed method use different values of size of neighbourhood, $a \in \{1,2,\cdots,10\}$, and choose the value of a that corresponds to the minimum value of RMSE, such that our final fused image

$$F = F_{a_0} \text{where} a_0 \in \{1,2,\cdots,10\} \text{ and } RMSE(F_{a_0}) = \min_{a\in\{1,2,\cdots,10\}} (RMSE(F_a))$$

5. EXPERIMENTAL RESULT

In this section, we are using images taken from the web page database [35]. We blurred an area of each reference image by applying the convolution of the Gaussian filter. The works [33] - [34] justify the choice of the Gaussian filter. The reference images contain at least two objects of the photographed scene. We have chosen to hide an object from the reference image. Thus from each image we obtain multi-focus images whose number is equal to the number of objects belonging to the reference image. The size of the blurred areas depends on the size of the masked object. We applied the approach on 150 images of the web page [35]. To make this article easy to read, we have chosen to present only three reference images that we blurred by masking an object each time to extract multi-focus images from each image (figures 4, 5, 7, 8, 10 and 11). Figures 6, 9 and 12 show the images fused using the proposed method. Visually, the image obtained by the proposed method gives a very satisfactory fusion.

Fig.4 in focus on the right

Fig.5 in focus on the left

Fig.6 Fused image by proposed method

Fig.7 in focus on the left

Fig.8 in focus on the right

Fig.9 Fused image by proposed method

Fig.10 in focus on the left

Fig.11 in focus on the right

Fig.12 Fused image by proposed method

To compare the proposed method to other methods, we perform the fusion using methods: PCA method [1], discrete wavelet transform (DWT) method [6], Laplacian pyramid LP_PCA [13], LP_DWT [14] and gradient bilateral (BG) [2].

To objectively evaluate these fusion methods, we will use quantitative measurements on the fused images. The RMSE evaluation measure was chosen for its efficiency in this comparison category. The table gives the mean and standard deviation of RMSE for the methods studied.

Table 1. Statistic parameters of the sample (150 images)

Method	LP_AV	PCA	BG	LP.PCA	DWT	LP.DWT	Proposed_method
Mean	6.351	6.245	7.7375	1.7456	3.0738	1.7841	0.44059
Standard deviation	2.81099	2.76977	3.77837	0.62897	1.06387	0.638727	0.223299

From the table1. We deduce that the proposed method has a smaller mean of the RMSE. The histograms of RMSE for 150 images by different methods (Figure 14, 15, 16, 17, 18 and 19) are for almost method symmetric and centred around the mean value.

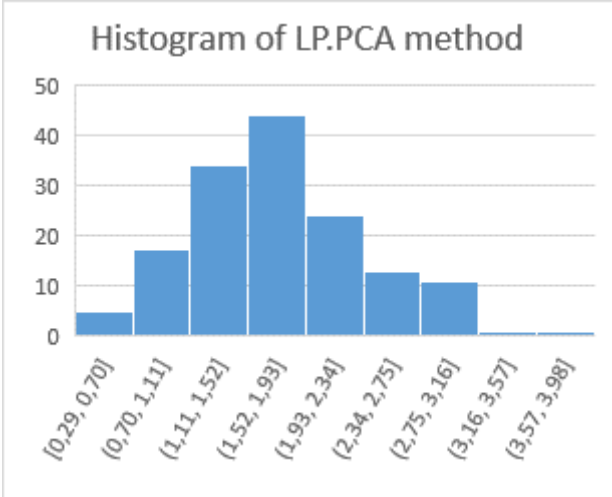

Figure 13. The histogram of LP. PCA method

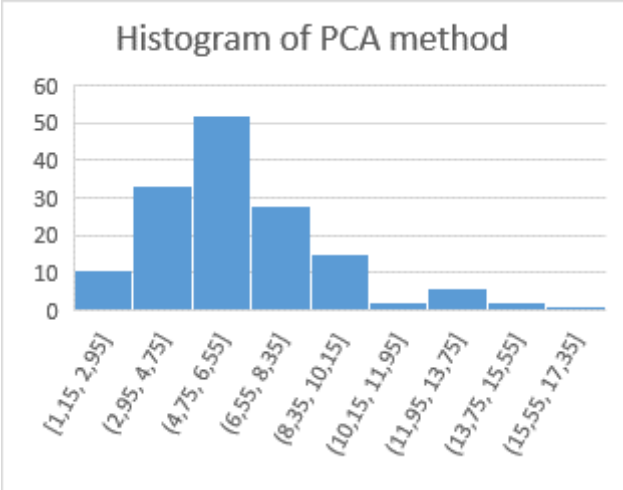

Figure 14. The histogram of PCA method

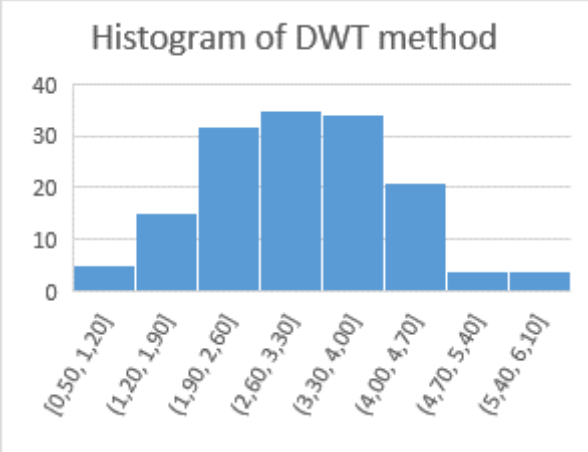

Figure 15. The histogram of DWT method

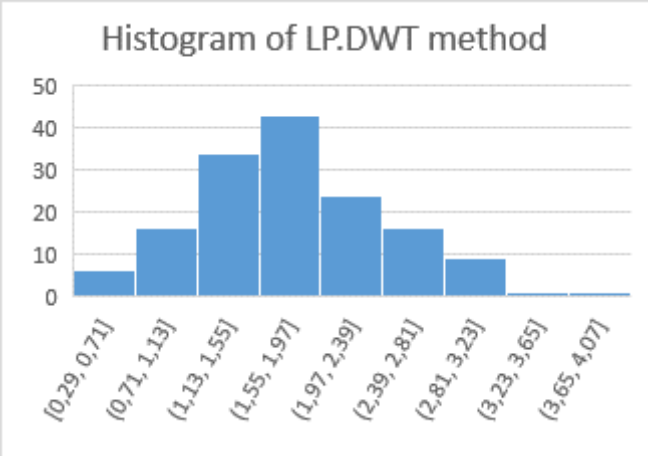

Figure 16. The histogram of LP.DWT method

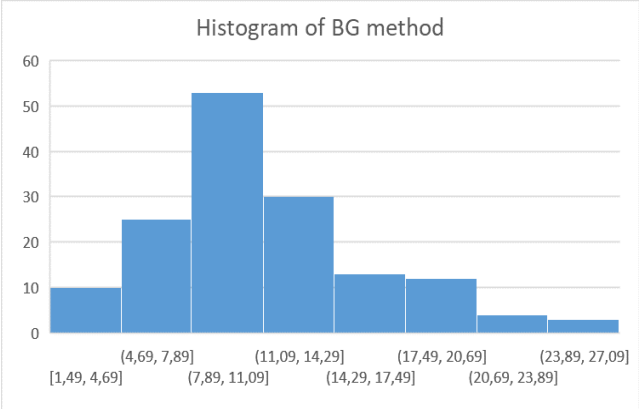

Figure 17. The histogram of BG method

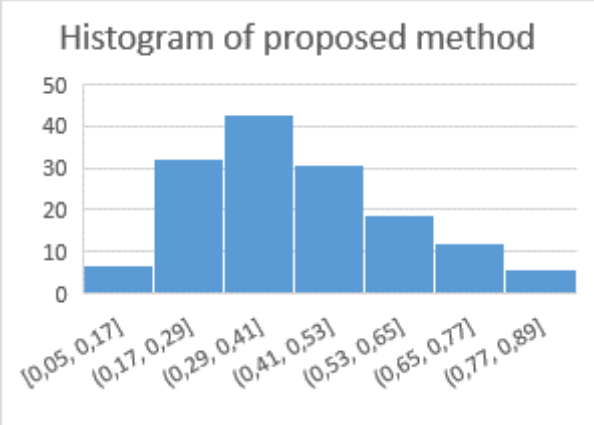

Figure 18. The histogram of proposed method

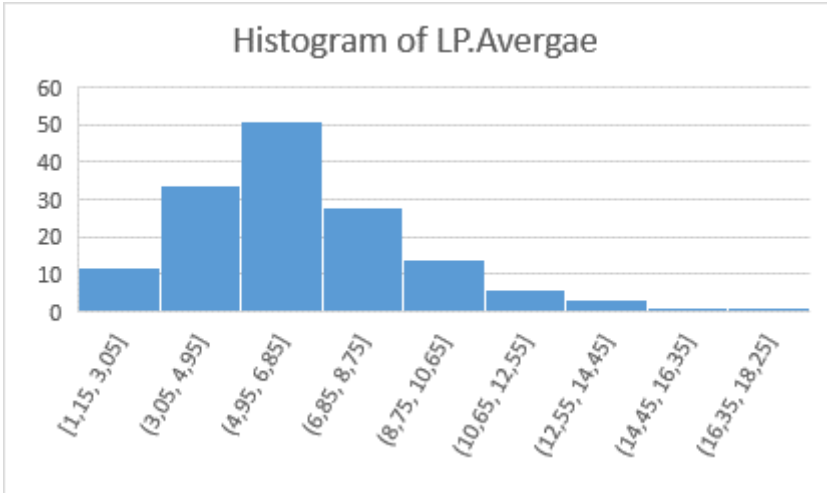

Figure 19. The histogram of LP.Average method

An analytic comparison, Analysis of variance (ANOVA) with dependent samples (dependence by image), is used. The software R gives the following Anova table:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Method	6	1811	301.76	65.92	<2e-16
Residuals	1036	4743	4.58		

As Pr(>F) is smaller than 1% the methods are significantly different. We use now the Newman Keuls test to compare the methods two-by-two and make groups having significantly the same mean. The software R gives the results below of the test.

```
$means
                    RMSE       std   r       Min       Max       Q25       Q50       Q75
BG             7.7374812 3.7783714 150 0.7996276 22.787395 5.4333421 6.9297521 9.8271991
DWT            3.0737953 1.0638734 150 0.5012112  6.030012 2.3205412 2.9734891 3.8514298
LP_AV          6.3513977 2.8109892 150 1.1543230 17.314372 4.6162405 5.9611743 7.5387607
LP_DWT         1.7841139 0.6387276 150 0.2922930  3.664108 1.3388044 1.7580606 2.1669267
LP_PCA         1.7456269 0.6289764 150 0.2871605  3.637775 1.3227422 1.7258679 2.1305223
PCA            6.2446737 2.7697695 150 1.1534551 17.155190 4.5865356 5.9243098 7.3428533
proposed_method 0.4405923 0.2232999 150 0.0549311  1.658533 0.2882735 0.3993412 0.5431666

$comparison
NULL

$groups
                    RMSE groups
BG             7.7374812      a
LP_AV          6.3513977      b
PCA            6.2446737      b
DWT            3.0737953      c
LP_DWT         1.7841139      d
LP_PCA         1.7456269      d
proposed_method 0.4405923      e
```

We obtain Four different groups: First Group “a” contains only method BG having the bigger mean of RMSE (7.737). The Group “b” contains two methods LP_AV and PCA that have significantly the same average. Group “c” contains only the method DWT which better than group “a” and “b”. Group “d” contains two methods LP_DWT and LP_PCA which better than

group “a”, “b” and “c”. The last group “e” containing the proposed method that the best method because his mean is the smallest by comparing with other means.

6. CONCLUSION

This paper, presents a new method to fuse two multi-focus images based on Dempster-Shafer theory using local variability. The originality of this method lies in the fact that we use the Dempster Shapher theory and the local variability of each pixel according to the quadratic distance. The fusion decision is obtained by pixels, which correspond to the minimum of plausibility. The experimental study shows that the proposed method gives a significant improvement in the result both visually and quantitatively. We extend this method to fuse more than two blurred images. We can use it in many applications, such as

1. The drone is becoming an essential tool in digital imaging, it offers interesting possibilities to improve photography. The drone can capture images on the same scene, which zooms in on different objects and at different altitudes. Thus, it will give several images on the same scene but with different objects in focus. With the proposed method, we obtain an image with all the objects at the point very similar to the real images
2. The method can also be used in medical imaging. In fact, it can be used to detect an object or cell anomaly due to the local variability indicating the behaviour of each pixel with its neighbourhood.
3. The food industry uses cameras to control the quality of the manufactured product. Each camera targets one of several objects to detect an anomaly. We would use the proposed method to get a photo containing all the objects in focus with more detailed information.

This work has several perspectives:

1. Our proposed method is performed on images in gray levels. We plan to extend it to colour images that convey important information.
2. Encouraged by the fusion of two images we intend to extend the method to more than two images by taking into account the local variability in each image (intra variability) and the variability between the images (inter variability). This inter-variability can detect "abnormal pixels" among images.
3. We are motivated to use the proposed method to fuse images containing different objects from different sensors (multimodal).

7. ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their important suggestions and remarks.

REFERENCES

[1] Naidu V.P.S. and. Raol, J.R., (2008) “Pixel-level Image Fusion using Wavelets and Principal Component Analysis”, Defence Science Journal, Vol. 58, No. 3, pp. 338-352.
[2] Tian, J., Chen, L., Ma, L., and Yu, W., (2011) “Multi-focus image fusion using a bilateral gradient-based sharpness criterion”, Optic Communications, 284, pp 80-87.
[3] Zhan, K., Teng, J., Li, Q., and Shi, J., (2015) “A novel explicit multi-focus image fusion method”, Journal of Information Hiding and Multimedia Signal Processing, vol. 6, No. 3, pp.600-612.
[4] Mallat, S.G., (1989) “A Theory for multiresolution signal decomposition: The wavelet representation”, IEEE Trans. Pattern Anal. Mach. Intel., Vol. 11, No. 7, pp.674-93.

[5] Pajares, G., Cruz, J.M., (2004) "A Wavelet-Based Image Fusion Tutorial", Pattern Recognition 37. Science Direct.
[6] Guihong, Q., Dali, Z., and Pingfan, Y., (2001) "Medical image fusion by wavelet transform modulus maxima". Opt. Express 9, pp. 184-190.
[7] Indhumadhi, N., Padmavathi, G., (2011) "Enhanced Image Fusion Algorithm Using Laplacian Pyramid and Spatial Frequency Based Wavelet Algorithm", International Journal of Soft Computing and Engineering (IJSCE). ISSN: 2231-2307, Vol. 1, Issue 5,.
[8] Burt, P.J., Adelson, E.H., (1983) "The Laplacian Pyramid as a Compact Image Code", IEEE Transactions on communication, Vol.Com-31, No 40.
[9] Burt, P.J., (1984) "The Pyramid as a Structure for Efficient Computation", Multiresolution Image Processing and Analysis, A. Rosenfeld, Ed., Springer-Verlag. New York 1984.
[10] Burt, P.J., Kolezynski, R.J., (1993) "Enhanced Image Capture Through Fusion", International Conference on Computer Vision, pp. 173-182.
[11] Wang, W., Chang, F., (2011) "A Multi-focus Image Fusion Method Based on Laplacian Pyramid", Journal of Computers, Vol.6, No 12.
[12] Zhao, P., Liu, G., Hu, C., Hu, and Huang, H., (2013) "Medical image fusion algorithm on the Laplace-PCA". Proc. 2013 Chinese Intelligent Automation Conference, pp. 787-794.
[13] Verma, S. K., Kaur, M., and Kumar, R., (2016) "Hybrid image fusion algorithm using Laplacian Pyramid and PCA method", Proceeding of the Second International Conference on Information and Communication Technology for Competitive Strategies.
[14] Wahyuni, I. S. and Sabre, R., (2016) "Wavelet Decomposition in Laplacian Pyramid for Image Fusion", International Journal of Signal Processing Systems Vol. 4, No. 1. pp. 37-44 . 2016.
[15] Haghighat, M. B. A., Aghagolzadeh, A., and Seyedarabi, H., (2010) "Real-time fusion of multifocus images for visual sensor networks", Machine vision and image processing (MVIP), 2010 6th Iranian.
[16] Bavirisetti, D. P. and Dhuli, R., (2016) "Multi-focus image fusion using multi-scale image decomposition and saliency detection", Ain Shams Eng. J., to be published. [Online]. Available: http://dx.doi.org/10.1016/j.asej.2016.06.011.
[17] Yuan, X., Zhang, J., Yuan, X., and Buckles, B.P, 2003) "Low level fusion of imagery based on Dempster-Shafer theory", Proceedings of the Sixteenth International Florida Artificial Intelligence Research Society Conference, pp. 475-479.
[18] Denoeux, T., (1999) "Reasoning with imprecise belief structures", International Journal of Approximate Reasoning Vol. 20, No.1, pp.79-111.
[19] Walley, P.,(1991) Statistical reasoning with imprecise probabilities, London: Chapman and Hall.
[20] Mena, J.B. and Malpica, J.A., (2003) "Color image segmentation using the Dempster-Shafer theory of evidence for the fusion of texture", Proceeding ISPRS, Vol. XXXIV-3/W8, pp. 139-144.
[21] Rombaut,M. and Zhu, Y.M., (2002) "Study of Dempster-Shafer for image segmentation application", Image vision Computing, Vol. 20, pp. 15-23.
[22] Kowsalya, M. and Yamini, C., (2015) "A survey on pattern classification with missing data using Dempster-Shafer theory", International conference on information engineering, management, and security: 134-138.
[23] Zhu, H., Basir, O., and Karray, F., (2002) "Data fusion for pattern classification via the Dempster-Shafer evidence theory", Proceeding IEEE Conf. Syst,. Man, Cybern., vol. 7, pp. 109-114.
[24] Hassan, M. H., (1989) "Object recognition based on Dempster-Shafer reasoning", Proc. SPIE 1002, Intelligent robots and computer vision VII.
[25] Bloch, I., (1996) "Some aspects of Dempster-Shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account", Pattern Recognition Letters 17, pp. 905-919.
[26] Wu, H., Siegel, M., Stiefelhagen, R, and Yang, J., (2002) "Sensor fusion using Dempster-Shafer theory", IEEE Instrumentation and Measurement Technology Conference Anchorage, AK, USA, 21-23.
[27] Klir, G.J and Folger, T. A., (1988) Fuzzy sets, uncertainty and information, Englewood Cliffs Prentice-Hall.
[28] Shafer, G., (1976) A mathemaical theory of evidence. Pricenton University Press.
[29] Dempster, A. P., (1967) "Upper and lower probabilities induced by a multivalued mapping", The Annals of Mathematical Statistics Vol.38, pp. 325-339.
[30] Dempster, A. P., 1968) "A genralization of Bayesian inference", Journal of the Royal Statistical Society, Series B (methodological) Vol. 30, pp. 205-247.

[31] Bloch, I., (2008) Information fusion in signal and image processing. John Wiley and Sons, Inc.
[32] Wahyuni I. Sabre R., (2020) "Neighbour Local Variability for Multi-Focus Images Fusion", Signal & Image Processing: An International Journal (SIPIJ) Vol.11, No.6, pp. 37-51.
[33] Nayar, S. K., (1992) "Shape from Focus System", Proc. of IEEE Conf. Computer Vision and Pattern Recognition, pp. 302-308.
[34] Petland, A., (1987 "A new sense for depth of field.", IEEE Transactions on Pattern Analysis and Machine Intelligent, Vol. 9, No. 4, pp. 523-531.
[35] www.rawsamples.ch. Accesed: 30 March 2018.
[36] Tianjing Feng, Hairong Ma, Xinwen Cheng , (2021) "Land-cover classification of high-resolution remote sensing image based on multi-classifier fusion and the improved Dempster–Shafer evidence theory", Journal of Applied Remote Sensing Vol. 15, No. 1, https://doi.org/10.1117/1.JRS.15.014506
[37] Ling Huang, Thierry Denoeux, David Tonnelet, Pierre Decazes, Su Ruan, (2021) "Deep PET/CT fusion with Dempster-Shafer theory for lymphoma segmentation", Electrical Engineering and Systems Science, submitted on 11 Aug 2021

AUTHORS

Ias Sri Wahyuni was born in Jakarta, Indonesia, in 1986. She earned the B.Sc. and M.Sc. degrees in mathematics from the University of Indonesia, Depok, Indonesia, in 2008 and 2011, respectively. In 2009, she joined the Department of Informatics (COMPUTING) System, Gunadarma University, Depok, Indonesia, as a Lecturer. She is currently a PhD student at University of Burgundy, Dijon, France. Her current research interests include statistics and image processing.

Rachid Sabre received the PhD degree in statistics from the University of Rouen, Rouen, France, in 1993 and Habilitation (HdR) from the University of Burgundy, Dijon, France, in 2003.He joined Agrosup Dijon, Dijon, France, in 1995, where he is an Associate Professor. From 1998 through 2010, he served as a member of "Institut de Mathématiques de Bourgogne", France. He was a member of the Scientific Council AgroSup Dijon from 2009 to 2013. In 2012, he has been a member of "Laboratoire Electronique, Informatique, et Image" (Le2i), France. Since 2019 has been a member of Laboratory Biogeosciences UMR CNRS, University of Burgundy. He is author/co-author of numerous papers in scientific and technical journals and conference proceedings. His research interests lie in areas of statistical process and spectral analysis for signal and image processing.

COMBINING EVIDENCE FROM AUDITORY, INSTANTANEOUS FREQUENCY AND RANDOM FOREST FOR ANTI-NOISE SPEECH RECOGNITION

Kun Liao

China Power Complete Equipment Co., Ltd, China

ABSTRACT

Due to the shortcomings of acoustic feature parameters in speech signals, and the limitations of existing acoustic features in characterizing the integrity of the speech information, This paper proposes a method for speech recognition combining cochlear feature and random forest. Environmental noise can pose a threat to the stable operation of current speech recognition systems. It is therefore essential to develop robust systems that are able to identify speech under low signal-to-noise ratio. In this paper, we propose a method of speech recognition combining spectral subtraction, auditory and energy features extraction. This method first extract novel auditory features based on cochlear filter cepstral coefficients (CFCC) and instantaneous frequency (IF), i.e., CFCCIF. Spectral subtraction is then introduced into the front end of feature extraction, and the extracted feature is called enhanced auditory features (EAF). An energy feature Teager energy operator (TEO) is also extracted, the combination of them is known as a fusion feature. Linear discriminate analysis (LDA) is then applied to feature selection and optimization of the fusion feature. Finally, random forest (RF) is used as the classifier in a non-specific persons, isolated words, and small-vocabulary speech recognition system. On the Korean isolated words database, the proposed features (i.e., EAF) after fusion with Teager energy features have shown strong robustness in the nosiy situation. Our experiments show that the optimization feature achieved in a speech recognition task display a high recognition rate and excellent anti-noise performance.

KEYWORDS

Cochlear filter cepstral coefficients; Teager energy features; Linear discriminate analysis; Random forest; speech recognition.

1. INTRODUCTION

Speech is the material shell and acoustic representation of language, and is one of the most direct, common and convenient carrier of information exchange for humans, and plays an important role in human-computer interaction and information transmission. With the advent of artificial intelligence (AI), it has always been the ideal of AI researchers to enable computers to simulate human consciousness and thinking information, so as to achieve human-computer interaction [1]. Speech recognition is equivalent to "machine's auditory system". It takes speech signal as the research object, combines signal processing technology and pattern recognition model to communicate with a computer, so that the computer can convert the speech signal into corresponding text or commands through the process of understanding and recognition [2]. Interpretation of human spoken language through technology has a diverse range of applications including in air transport, intelligent homes, disaster rescue, medical diagnostics, and other

David C. Wyld et al. (Eds): CSTY, SIGI, MaVaS - 2021
pp. 75-85, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.112207

human-computer interaction fields [3]. In the aviation transmission industry, speech recognition technology is used in the monitoring and response of the aviation air traffic control command, which makes the control command automatically respond, and then improves the training efficiency of the analog aircraft [4]. In smart home, combining speech recognition technology, wireless information transmission technology and embedded mobile computing technology, we can design a barrier-free intelligent home environment system for the disabled groups, so as to help them improve their quality of life and ability to participate in society [5]. In disaster rescue, speech recognition technology is integrated into the design of sensors to detect the calls of the victims, realize the auditory navigation of rescue robots, and achieve the purpose of search and rescue [6]. In medical diagnosis, according to the relationship between voice and depression, a depression recognition model based on voice variables is built to realize the correct diagnosis of the disease [7]. All the above studies are the practical application of speech recognition technology in social life, which shows that speech recognition technology has high research value and significance, and the improvement of the performance of speech recognition system has become a research hotspot of researchers.

As an important element of speech recognition, feature extraction has a large influence on the performance of the system [8]. Therefore, methods to extract the most information-capable, noise-less, easily classified and stable new features from speech must be developed, and integrate and optimize the different types of features that have been proposed also require further research to establish a speech recognition system with the best classification performance. Currently, the most mainstream speech feature is Mel frequency cepstral coefficient (MFCC), and MFCC is extracted based on Fourier transform, studies have shown that the Fourier transform is not suitable for the processing of non-stationary time-varying signals [9]. In recent years, a new method of auditory transformation compensates for the shortcomings of Fourier transform. The time-frequency analysis method of introducing auditory transformation into speech signal processing has attracted the attention of some researchers, and based on auditory transformation, a new feature model which is more in line with human ear auditory characteristics has been proposed. For instance, Peter Li try to extract cochlear filter cepstral coefficients (CFCC) features for robust speaker identification [10-11], and CFCC features show strong robustness in speech signal processing [12]. In [13], Yanyan proposed CFCCIF features based on Cochlear filter cepstral coefficients and instantaneous frequency, and combined CFCCIF features with Principal Component Analysis (PCA) for speech recognition. Although there are many applications of CFCC features, there are very few studies focused on the application of CFCCIF features to speech recognition in noisy environments. Hence, we attempt to apply CFCCIF features to noisy speech recognition systems. Spectral subtraction is then introduced in the front-end of CFCCIF features extraction to enhance features, so as to extract more robust feature parameters. And we combine enhanced CFCCIF and Teager energy features to form a fusion feature, and linear discriminate analysis (LDA) is used to optimize the fusion feature parameters to obtain the optimal feature to improve the recognition accuracy. Finally, random forest (RF) is used as the classifier in a speech recognition system.

2. PROPOSED CFCCIF FEATURES

2.1. Cochlear filter cepstral coefficients (CFCC)

The feature extraction procedure for CFCC consists of four parts: a series of cochlear filter banks model based on auditory transform, hair cell function, nonlinearity, and discrete cosine transform (DCT)[14]. The following subsection briefly describes auditory transform and procedure for estimating the CFCC features.

2.1.1. Auditory Transform

As a new method of processing non-linear signals, auditory transform is equivalent to converting time-domain signals into frequency-domain signals through cochlear filter banks. The cochlear filter function is used as the basis function of the wavelet, completing the whole process of sound transmission from the outer ear to the basement membrane, with an existing inverse transform [15].

Let $\boldsymbol{\psi(t)}$ be the impulse response of the basilar membrane of cochlear $\psi(t) \in L^2(R)$, in which the function $\psi(t)$ satisfies the following conditions:

①. It integrates to zero:

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \tag{1}$$

②. It is square integrable or has finite energy:

$$\int_{-\infty}^{+\infty} \left| \psi(t) dt \right|^2 < \infty \tag{2}$$

③. It satisfies:

$$\int_{-\infty}^{+\infty} \frac{\left|\Psi(\omega)\right|^2}{\omega} d\omega = C \tag{3}$$

where $0 < C < \infty$, and

$$\Psi(\omega) = \int_{-\infty}^{+\infty} \psi(t) e^{-j\omega t} d\omega \tag{4}$$

Let $f(t)$ be any square integrable function. The auditory transform of $f(t)$, with respect to $\psi(t)$ as the impulse response of the basilar membrane in the cochlea, is defined as:

$$T(a,b) = \int_{-\infty}^{+\infty} f(t) \psi_{a,b}(t) dt \tag{5}$$

where $\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) dt$, a, b are real, and a is a scale or dilation variable. By changing a, the central frequency of an impulse response function can be shifted. Subscript b is a time shift or translation variable. If a is known, $\psi_{a,0}(t)$ moves a unit along the time axis to get $\psi_{a,b}(t)$. Note that $1/\sqrt{a}$ is an energy normalizing factor. It ensures that the energy stays the same for all a and b, providing:

$$\int_{-\infty}^{+\infty} \left|\psi_{a,b}(t)\right|^2 dt = \int_{-\infty}^{+\infty} \left|\psi(t)\right|^2 dt \tag{6}$$

A typical cochlear impulse response function or cochlear filter can be defined as:

$$\psi_{a,b}(t)=\frac{1}{\sqrt{a}}\frac{(t\text{-}b)^{\alpha}}{a}\exp\left[-2\pi f_L\beta\left(\frac{t\text{-}b}{a}\right)\right]\times\left[\cos 2\pi f_L\left(\frac{t\text{-}b}{a}\right)+\theta\right]u(t\text{-}b) \tag{7}$$

where $\alpha>0,\beta>0$, parameters α and β determine the frequency domain shape and width of the cochlear filter. Subscript α and β are taken as the generally empirical value, $\alpha=3,\beta=0.2$, $u(t)$ is the unit step function, and the value θ is the initial phase. The value of a can be determined by the current filter, the central frequency f_c, and the lowest frequency f_L of the cochlear filterbank, which is denoted as:

$$a=f_L/f_c \tag{8}$$

2.1.2. Other Operations in CFCC Extraction

As an important part of the auditory system, human cochlear inner ear hair cells transform the vibration signals transmitted from the basement membrane of the cochlea into analyzable nerve impulse signals of the brain, and then transmit them to the auditory nerve fibers[16]. The following nonlinear function of hair cell describes this motion:

$$h(a,b)=[T(a,b)]^2 \tag{9}$$

where $T(a,b)$ is the filterbank output of speech signal $f(t)$.

The hair cell output of each filterbank is converted into a representation of the nerve spike count density in a duration associated with the current band central frequency, which is computed as:

$$S(i,j)=\frac{1}{d}\sum_{b=1}^{l+d\text{-}1}h(i,b),l=1,L,2L\cdots;\forall i,j \tag{10}$$

where $d=max\{3.5\tau_i,20ms\}$ is the window length, τ_i is the period of the i band. $\tau_i=1/f_c$, and L is the window shift duration.

The output of the above formula is further applied to scales of loudness functions as logarithmic nonlinearity, providing:

$$y(i,j)=\log\left[S(i,j)\right] \tag{11}$$

Finally, discrete cosine transform (DCT) is applied to decorrelate the feature dimensions. It generates the cochlear filter cepstral coefficients as a new auditory-based speech feature, which is computed as:

$$cfcc(i,n)=\sqrt{2/M}\sum_{m=1}^{M-1}y(i,m)\cos\left(\frac{\pi n(m-1/2)}{M}\right),\quad 0\le m\le M \tag{12}$$

where M is the number of filters.

2.2. Instantaneous frequency (IF) estimation

Let $s(t)$ be speech signal, the IF of $s(t)$ is defined as the derivative of the unwrapped phase of the analytic signal derived from $s(t)$. For a real signal $s(t)$, its complex analytic representation is given by:

$$s_a(t) = s(t) + js_h(t) \tag{13}$$

where $s_h(t)$ is the Hilbert transform of $s(t)$, and the inverse Fourier transform(IFT) of $s_h(t)$ can be expressed as:

$$S_h(\omega) = \begin{cases} +jS(\omega) & \omega < 0 \\ -jS(\omega) & \omega > 0 \end{cases} \tag{14}$$

Then, the amplitude envelope of $s_a(t)$ is expressed as:

$$\left| s_a(t) \right| = \sqrt{s^2(t) + s_h^2(t)} \tag{15}$$

The instantaneous phase is $\phi(t) = \tan^{-1}\left(s_h(t)/s(t)\right)$, and IF is derived from unwrapped instantaneous phase, which can be expressed as:

$$IF = \frac{d}{dt}\left(\phi(t)\right) \tag{16}$$

2.3. CFCCIF Estimation

Similar to nerve spike density estimation, the IF is obtained as:

$$SIF(i,\mathrm{j}) = \frac{1}{d}\sum_{b=l}^{l+d-1} IF\left(h(i,b)\right), \qquad l = 1, L, 2L, \cdots; \forall i, j \tag{17}$$

To use both envelope structure and IF information, the IF features (eq.15) are multiplied with the corresponding nerve spike density envelope (eq.10). Thus, IF obtained in silence regions will be suppressed. To capture the transient information, the change in envelope and IF between consecutive frames is estimated through derivative operation followed by logarithm. Finally, DCT is applied framewise to get CFCCIF features.

3. PROPOSED EAF FEATURES

3.1. Spectral Subtraction

The principle of spectral subtraction is that the power spectrum of pure speech signal can be obtained by subtracting the power spectrum of noise from the power spectrum of speech signal with noise [17].

Let $y(n)$ be speech signal with noise, $s(n)$ is pure speech signal, $d(n)$ is noise, and the relationship between them is:

$$y(n)=s(n)+d(n),0\le n\le N\text{-}1 \tag{18}$$

where n the data points, and N is frame length.

Their representation in the Fourier transform domain is given by:

$$Y(\omega)=S(\omega)+D(\omega) \tag{19}$$

As speech is assumed to be uncorrelated with background noise, the short-term power spectrum of $y(n)$ has no cross-terms, hence:

$$E\left|\left|S(\omega)\right|^2\right|=E\left|\left|Y(\omega)\right|^2\right|-E\left|\left|D(\omega)\right|^2\right| \tag{20}$$

where $S(\omega)$, $D(\omega)$, $Y(\omega)$ is the short-term power spectrum of $s(n)$, $d(n)$, and $y(n)$.

For a short-time stationary process in a frame, use:

$$\left|S(\omega)\right|^2=\left|Y(\omega)\right|^2-\lambda_d(\omega) \tag{21}$$

in which $\lambda_d(\omega)$ is the statistical average of silent segment $\left|D(\omega)\right|^2$. Therefore, the amplitude of the speech signal after spectral subtraction can be expressed as:

$$\hat{S}(\omega)=[|Y(\omega)|^2\text{-}E(|D(\omega)|^2)]^{1/2}=[|Y(\omega)|^2\text{-}\lambda_d(\omega)]^{1/2} \tag{22}$$

3.2. EAF Features Extraction

To reduce the influence of noise on the CFCCIF features and further enhance the robustness of features, spectral subtraction is introduced in the front-end of CFCCIF features extraction. The speech signal is preprocessed first in a process which includes pre-emphasis, endpoint detection, and frame windowing. Formula (20) is then used to subtract the spectrum amplitude of noise from the spectrum amplitude of the noise signal, providing the spectrum amplitude of pure signal. Based on the phase insensitivity of speech, the phase angle information before spectral subtraction is directly used to reconstruct the signal after spectral subtraction to obtain the denoised speech. Finally, the denoised speech signal is extracted using process of CFCCIF features extraction, the enhanced auditory features (EAF)is obtained. The extraction process is illustrated in Figure. 1.

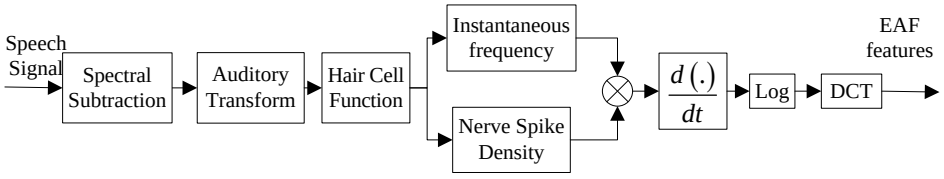

Figure 1. Extraction process

4. FUSION FEATURES EXTRACTION AND OPTIMIZATION

4.1. Teager Energy Features

Let $x(n)$ be a discrete-time signal, and the definition of TEO is:

$$\psi[x(n)] = x(n)^2 - x(n+1)x(n-1) \tag{23}$$

where $\psi[x(n)]$ is output of TEO, $x(n)$ is the sampling value of the discrete signal at n point.

Let $x(n)$ be a speech signal with additive noise, $s(n)$ be pure speech signal, and $\omega(n)$ be zero-mean additive noise. This relationship can be expressed as:

$$x(n) = s(n) + \omega(n) \tag{24}$$

The TEF of $x(n)$ is given by:

$$\psi[x(n)] = \psi[s(n)] + \psi[\omega(n)] + 2\tilde{\psi}[s(n), \omega(n)] \tag{25}$$

where $\tilde{\psi}[s(n), \omega(n)]$ is mutual Teager energy of $s(n)$ and $\omega(n)$, and

$$\tilde{\psi}[s(n), \omega(n)] = s(n)\omega(n) - 0.5s(n-1)\omega(n+1) - 0.5s(n+1)\omega(n-1) \tag{26}$$

Both $s(n)$ and $\omega(n)$ are zero mean and independent of each other, providing:

$$E\{\tilde{\psi}[s(n), \omega(n)]\} = 0 \tag{27}$$

$$E\{\psi[x(n)]\} = E\{\psi[s(n)]\} + E\{\psi[\omega(n)]\} \tag{28}$$

Compared with TEF of pure speech signal, the TEF of noise can be neglected, according to:

$$E\{\psi[x(n)]\} \approx E\{\psi[s(n)]\} \tag{29}$$

Thus, TEF can eliminate the influence of zero-mean noise and achieve speech enhancement [17-17]. The application of TEF in feature extraction can not only better track the non-linear energy of speech signal, can reasonably present the transformation of signal energy, but also suppress noise and enhance speech signals, achieving good results in speech recognition.

To construct a more effective subset of speech features, this paper combines above EAF features and TEF which reflects the change of signal energy. The fusion features not only represent the auditory perception characteristics and instantaneous frequency information of human ears, but also combine the characteristics of speech energy change, and suppress the zero-mean noise effect on speech signal to some extent, so as to more accurately describe the characteristics of speech.

4.2. Linear Discriminant Analysis (LDA)

To reduce the storage of feature data and further optimize the fusion feature, LDA is used to reduce the dimension of the fusion feature and further improve the performance of the recognition system. The principle of LDA is to project the high-dimensional pattern samples into the optimal discriminant vector space to extract classification information and compress the dimension of feature space. After projection, the pattern samples can have the largest class distance and the smallest intra-class distance in the new subspace, that is, the pattern has the best separability in the space [19-20]. It can transform the original feature set into a new feature subspace with a lower dimension, and compress the data while keeping as much relevant information as possible.

5. EXPERIMENTAL SETUP AND ANALYSIS OF RESULTS

The isolated words database is used for performing isolated word recognition from speech signals. The vocabulary sizes used here are 10 words and 20 words. The corpus consists of 10 digits and 40 command words. Random forest is used for speech recognition comparison experiments, and 10-fold cross-validation method is used to test the performance of feature and recognition network.

To verify the validity and robustness of the proposed fusion feature and optimization feature, and the optimized feature is defined as LDA-Features, the following experimental schemes are designed.

Table 1 Comparison of speech recognition based on five features (%)

Vocabulary	Experiment	Features	SNR (dB)					Average
			0	5	10	15	20	
10	Experiment 1	CFCCIF	80.79	86.63	87.12	90.41	90.96	87.18
	Experiment 2	EAF	82.94	87.38	89.08	91.00	91.80	88.44
	Experiment 3	EAF+TEF	83.48	88.02	89.08	91.80	92.06	88.89
	Experiment 4	LDA-Features	92.40	94.41	96.71	96.32	96.83	95.33
20	Experiment 1	CFCCIF	72.76	81.17	85.68	87.61	88.89	83.22
	Experiment 2	EAF	73.16	81.99	86.81	88.68	90.37	84.20
	Experiment 3	EAF+TEF	77.13	82.00	87.28	89.22	90.77	85.28
	Experiment 4	LDA-Features	90.30	94.20	94.27	96.05	95.55	94.07

(1) According to the recognition results of experiment 1 and experiment 2 shown in TABLE I, it can be observed that EAF features have a superior recognition effect compared with CFCCIF features. The average speech recognition rate under 10 words is increased by 1.26%, and under 20 words is increased by 0.98 %. This result proves that the feasibility and validity of the proposed features EAF in isolated word speech recognition systems. However, it can also been seen in TABLE I that the recognition effect of the CFCCIF features in a low SNR environment is not ideal.

(2) Comparing the recognition results of experiment 2 and experiment 3, it can be seen that after adding the TEF feature that embodies nonlinear energy characteristics, the recognition effect of the fusion feature is further improved compared to single feature. This result illustrates that TEF contains the effective information of speech signal and can be used as an auxiliary feature parameter to improve the performance of a speech recognition system.

(3) Comparison of experiment 2 and 3 confirms that after the fusion feature is optimized by LDA, the recognition rate have a significant improvement. This is because LDA can reduce dimension of feature while retaining important information in feature, thus improving the classification performance of speech recognition systems, and further verifying the effectiveness of fusion feature optimization.

(4) Figure. 2. depicts the speech recognition results of four groups of experiments in different SNR environments intuitively. It can be determined that fusion feature have certain advantages in recognition rate and robustness. It further illustrates the fusion feature can be constructed by combining auditory features with energy feature, and further verifying the effectiveness of LDA optimized feature method.

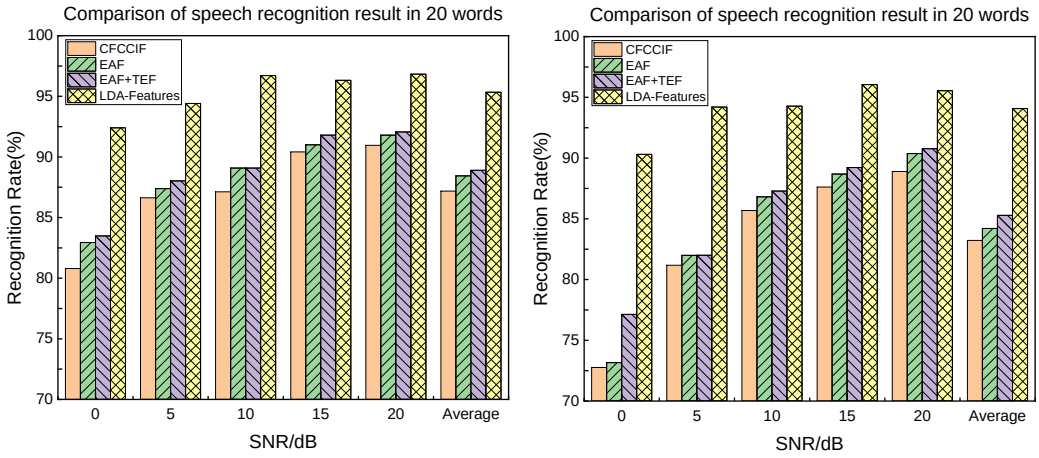

Figure. 2. Comparison of experimental results

6. CONCLUSION AND FURTHER STUDY

In this paper, CFCCIF features were extracted based on CFCC and instantaneous frequency information, and this feature was applied to speech recognition system in noisy environment. Spectral subtraction was then introduced into the front end of CFCCIF features extraction to improve the robustness of the feature to noise, and the extracted EAF features and TEF was combined in a fusion feature. It is proved that the fusion feature can effectively improve the recognition rate of speech recognition compared with the single feature. Finally, a feature optimization method of linear discriminate analysis was proposed, and its effectiveness was verified. In future research, we would like to consider finding a better speech enhancement method combined with feature extraction to achieve better speech recognition performance. In addition, the study of more better feature optimization are also the future research direction that needs to be further explored.

REFERENCES

[1] Ding I J, Yen C T . Enhancing GMM speaker identification by incorporating SVM speaker verification for intelligent web-based speech applications[J]. Multimedia Tools & Applications, 2015, 74 (14):5131-5140.

[2] Kłosowski, Piotr, Dustor A . Automatic Speech Segmentation for Automatic Speech Translation[M]// Computer Networks. Springer Berlin Heidelberg, 2013.

[3] Karpov A A, Yusupov R M . Multimodal Interfaces of Human–Computer Interaction[J]. Herald of the Russian Academy of Sciences, 2018, 88 (1):67-74.

[4] Cheung B . The resurgence of tactile display technologies.[J]. Aviation Space & Environmental Medicine, 2004, 75 (10):925.

[5] Han Y, Hyun J, Jeong T, et al. A smart home control system based on context and human speech[C]// 2016 18th International Conference on Advanced Communication Technology (ICACT). IEEE, 2016.

[6] Doostdar M, Schiffer S, Lakemeyer G . A robust speech recognition system for service-robotics applications.[J]. Lecture Notes in Computer Science, 2008, 5399:1-12.

[7] Lang H, Cui C . Automated Depression Analysis Using Convolutional Neural Networks from Speech[J]. Journal of Biomedical Informatics, 2018:S153204641830090X-.

[8] Wang G Y, Zhang Y M, Sun M L, et al. Speech signal feature parameters extraction algorithm based on PCNN for isolated word recognition[C]// International Conference on Audio. IEEE, 2017.

[9] Seyedin S, Gazor S, Ahadi S M . On the distribution of Mel-filtered log-spectrum of speech in additive noise[J]. Speech Communication, 2015, 67:8-25.

[10] Li Q L Q . An auditory-based transfrom for audio signal processing[C]// IEEE Workshop on Applications of Signal Processing to Audio & Acoustics. IEEE, 2009.

[11] Li Q, Huang Y. Robust speaker identification using an auditory-based feature[J]. IEEE International Conference on Acoustics Speech & Signal Processing, 2010, 23 (3):4514-4517.

[12] Li Z, Gao Y . Acoustic feature extraction method for robust speaker identification[M]. Kluwer Academic Publishers, 2016. 12 (75): 7391-7406.

[13] Yanyan, Shi, Jing, et al. Fusion Feature Extraction Based on Auditory and Energy for Noise-Robust Speech Recognition[J]. IEEE Access, 2019, 7:81911-81922.

[14] Wang D, Hansen J H L . F0 estimation for noisy speech by exploring temporal harmonic structures in local time frequency spectrum segment[C]// The 41st IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2016. 6510-6514.

[15] Li Q, Huang Y. An Auditory-Based Feature Extraction Algorithm for Robust Speaker Identification Under Mismatched Conditions[J]. IEEE Transactions on Audio Speech & Language Processing, 2011, 19 (6):1791-1801.

[16] Haji T, Kitazawa S . Acoustic analysis of certain consonants using a computed model of the peripheral auditory system.[J]. Nippon Jibiinkoka Gakkai Kaiho, 1994, 97 (11):2055-2064.

[17] López-Oller, Domingo, Benamirouche N, Gomez A M, et al. Speech excitation signal recovering based on a novel error mitigation scheme under erasure channel conditions[J]. Speech Communication, 2018, 97:73-80.

[18] Kaiser J F. On a simple algorithm to calculate the 'energy' of a signal[C]// International Conference on Acoustics, Speech, and Signal Processing. IEEE, 2002:381-384 vol.1.

[19] Senior A, Cho Y, Weston J . Learning improved linear transforms for speech recognition[C]// IEEE International Conference on Acoustics. IEEE, 2012.

[20] Gharsellaoui S, Selouani S A, Dahmane A O . Automatic emotion recognition using auditory and prosodic indicative features[C]// Electrical & Computer Engineering. IEEE, 2015.

AUTHORS

Short Biography

Kun Liao, Graduated from Northeast Electric Power University, research on Pattern Recognition, Artificial intelligence, image processing and speech recognition.

RESEARCH ON THE MODEL OF INTERACTION AND COUPLING RELATIONSHIP BETWEEN STANDARD SYSTEM CONSTRUCTION AND PROCESS MANAGEMENT

Wang Chaofan[1] and Chen Xinyue[2]

[1]North China Electric Power University, School of Economics and Management, Wang Chaofan, Beijing, China
[2]Qingdao University, School of Quality and Standardization, Chen Xinyue, Shandong, China

ABSTRACT

The fundamental motivation for enterprises to build a standard system is to meet the subjective needs of the unique use nature of standards, such as benchmarking and criteria, from input to output in the baseline relationship of all their business processes. How to guess the credibility and value of the standard system requires collaborative and mature processes to mediate cognition. Firstly, this paper clarifies the philosophical relationship between standard system construction and process management by using the general system structure theory. Secondly, it systematically summarizes the interaction mechanism between the two from the perspective of methodology. Finally, it designs a conceptual model with the "human regulation" composite system as the core and the coupling between standard system construction and the outer edge of process management, and combed the three mechanism characteristics that affect the fit of the coupling model. In order to provide a new integration idea for enterprise standardization management and process management to jointly realize the optimal value utility.

KEYWORDS

Standardization Discipline, Construction of Standard system, Process Management, Coupling Model.

1. INTRODUCTION

The fundamental motivation for an enterprise to build a standard system is to meet the subjective requirements for the unique use of standards, such as benchmarking and standardization, from the input end to the output end in the baseline relationship of all its business processes. The traditional standardization system is structured based on the function as the core, that is, through the in-depth exploration of the manager's management experience and the staff's work experience, the establishment of a corporate standardization system provides a systematic standardization management method and management platform for the realization of corporate strategies (Mai, 2018). This design method will inevitably be affected by the subjective role consciousness of the experience provider and the ideology of standardism, which strengthens the autonomy of the experience provider. It is unavoidable that some experience providers seek to maximize their personal interests or demonstrate their personal indispensability (Wang & Zhao, 2021). Conceal the individual's unique experience that should be accurately and unreservedly fed back to the standard-setting department as a "secret", or secretly lower the explicit truth threshold

David C. Wyld et al. (Eds): CSTY, SIGI, MaVaS - 2021
pp. 87-99, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.112208

of standard items that matches the job ability to invisibly encroach on job resources. This subjective behavior of the experience provider deviates from the original intention of the company to precipitate and reproduce its "skills" when constructing the standard system. At the same time, because internal control is essentially dependent on people, it is difficult to detect and control the scope of the "secret". So that no amount of PDCA cycles in the standardization work will eventually become ineffective consumption. At the same time, if the experience provider suddenly resigns and the successor performs the task with reference to its solidified experience standards, it will be difficult to implement the guidance of the standard, fail to achieve the expected work efficiency, produce repetitive technical labor and knowledge labor, and damage the existing the value utility of the standard system (Krislov & Samuel, 2020).

How to estimate the credibility and value of the standard system, reduce the risk of enterprise employment, and reduce the degree of dependence on special talents? This requires the introduction of the concept of process management. Process management is a systematic method centered on the standardized structure of end-to-end excellent business processes, and continuously improving the business performance of the organization as the value goal, emphasizing that the content, methods, and responsibilities of the process should be strictly prioritized and statistically controlled. The state is only affected by random factors (Harm Schepel, 2017). Only by controlling the input and process, the quality of the output can be guaranteed, which dilutes the subjectivity and spontaneity of people at work.

In essence, the enterprise standard system is a scientific and organic whole formed by summarizing various standards formulated and implemented around various business processes such as production, operation and management according to their relevance and internal relations (Jaesun Wang & Seoyong Kim, 2016). Therefore, the effective operation of the enterprise standard system depends on the normal operation of various business processes. This article intends to use general system structure theory to clarify the philosophical relationship between standard system construction and process management, and try to systematically summarize the interaction mechanism between the two from the perspective of methodology. Secondly, the design is based on the "human-regulation" composite system as the core, standard system construction and process. The conceptual model of external coupling of management is expected to provide new integration ideas for the joint realization of enterprise standardization management and process management to achieve the best value and utility.

2. THE PHILOSOPHICAL RELATIONSHIP BETWEEN STANDARD SYSTEM CONSTRUCTION AND PROCESS MANAGEMENT

A feasible and effective model must be based on the correct philosophical relationship. The correct philosophical relationship between the construction of the standard system and the process management should be recognized as: the construction of the standard system provides a theoretical basis for the effective operation of the process. The standard organization behavioral cognition view provides the source of feedback reform (Smith G, 2015). Go back to GB/T 20000.1's definition of standardization: in order to obtain the best order within the established scope, promote common benefits, establish common use and repeated use clauses, and compile, publish and apply documents for actual or potential problems. At the same time, as mentioned in the previous article, process management is based on the guidance of the logical thinking mode of the enterprise and the existing resource capabilities, and the control of the interconnected or interactive activities that transform input into output (Lawrence Busch, 2011).

From the comparison and analysis of the relationship between the generation and application of the two, the purpose of the construction of the standard system is to provide evidence-based

guarantees for the consistency of the continuity and quality of the production or service provided by the company. It is not an original factor for the company to generate value-added benefits, but it is an indispensable enhancement factor or improvement factor. Its benefit contribution is attributed to the multi-dimensional influence of standardization in improving enterprise technology, work, management, and hardware. At the same time, it can basically only evaluate the input and output factors related to benefit through the black box method. In addition, the total benefit of the construction of the standard system is the result of the full life cycle time integration of the standardization work; the purpose of process management is to improve the compliance and operability of all management objects in the enterprise that have the ability to execute unified behavior instructions, and the transmission of their behavior directly affects the operation results. Therefore, it is the original factor for the enterprise to generate value-added benefits. It is necessary to carry out internal restraint control and benefit evaluation in a white box manner to ensure that the results of the same behavior awareness are unified within the specified tolerance, and the total benefit is equal to the company's production of products or services. The overall value generated. From this analysis, it is concluded that the construction of the standard system and the process management both contribute to the benefits of the enterprise, and the two overlap in the affected objects. They complement each other in the evaluation of the benefits of the enterprise, and they are in a mutually dependent relationship.

Analyzed from the logical relationship between the outlines of the two expressing corporate management concepts, the construction of a standard system is the "donor" of standardization, and it is a process of subjectively studying standardization theories to arrive at a universally applicable method. The focus is on achieving excellent production with a unified and recognized characteristic relationship. The solidification of formal channels and procedures, the standard system is the standard sub-cluster, which is the objective relationship of enterprise management established based on the subjective recognition of the standardization discipline theory. Once the subjective establishment process is completed, it becomes an objective relationship that cannot be changed at will. The rigid rules consciously abided by within the enterprise are the "theoretical" representatives of the enterprise management philosophy; the process management is the "acceptor" of standardization, which is the objective application of the "benchmark" compliance relationship between the standardized objects in the input and output process. In the process of testing and inspection, a set of processes has a variety of standardized implementation methods and procedures. When carrying out process management, each manager will independently select and sequence the originally independent and uncoordinated sequence activities according to his own experience and background, so that the process management is flexible in operation and is the "practical" representative of corporate management concepts. The outline of the standard system construction is the standard system table, and the outline of the process management is the process operation manual. From this analysis, it is concluded that the standard system construction and process management are the standardized "active side" and the standardized "passive side" respectively, which are opposite to each other.

In summary, it is established that the construction of the standard system and the process management are mutually dependent and opposed to the unity of philosophy (Figure 1).

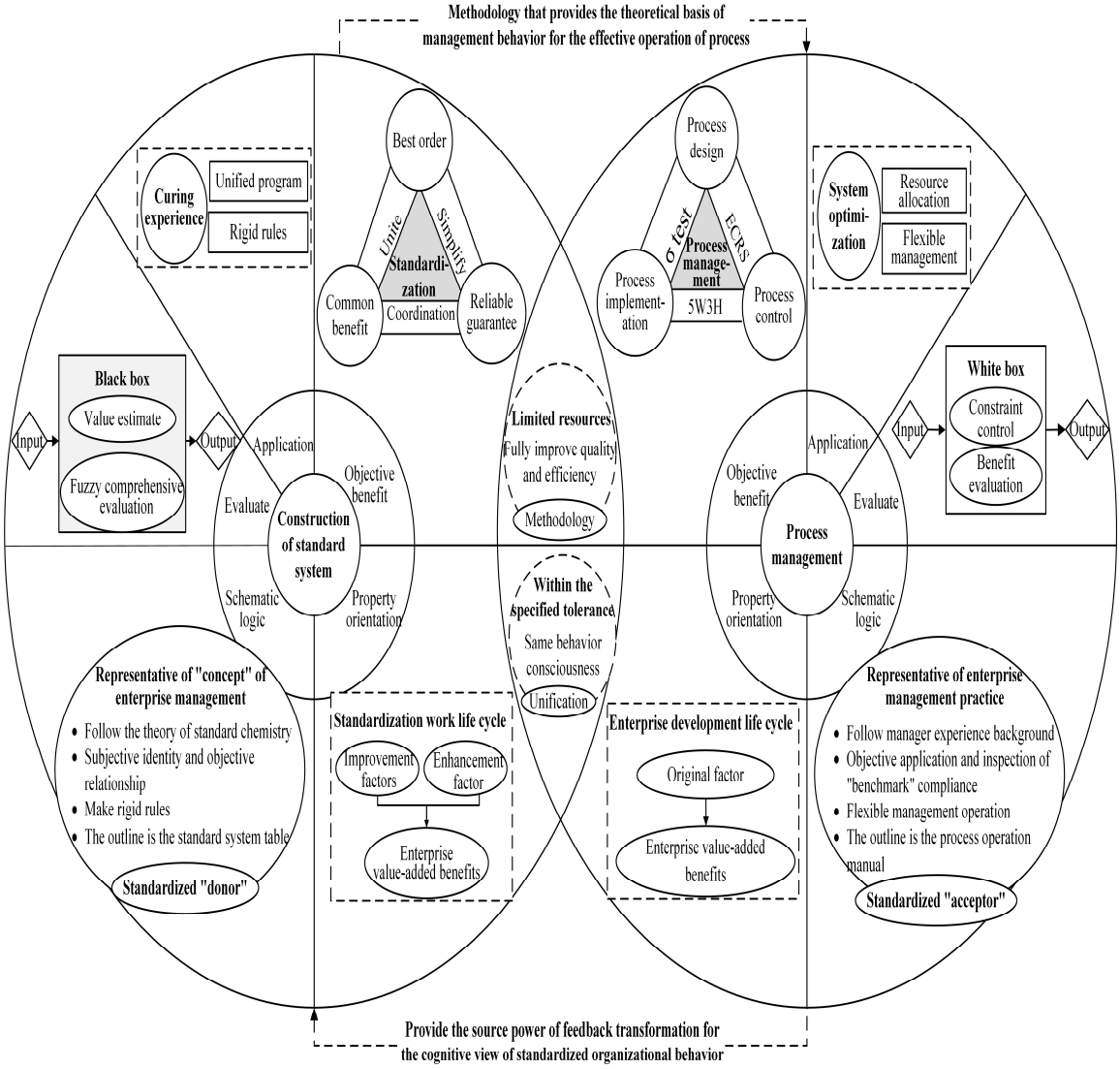

Figure 1. Philosophical relationship between standard system construction and process management.

3. THE INTERACTION MECHANISM OF STANDARD SYSTEM CONSTRUCTION AND PROCESS MANAGEMENT

3.1. The impact of standard system construction on process management

To know the skill is to be round, but to do it straight. Summarizing the role of standard system construction on process management, we might as well express the role of standardization on process management. At the vertical level, because the construction of the standard system is the necessary way and procedure for the implementation of the standardized method of process management, its essence is a standardized application process; in the horizontal scope, the purpose of the construction of the standard system is to organically integrate the standards of the various processes within the enterprise together, to achieve the coordination and unification of process standards and scientific integration, the ultimate goal is to establish the best order of process management for standardization and obtain the "absolute core" service with the best benefits. In summary, the construction of a standard system and standardization have the same

vertical measurement concept in summarizing the effect of standardization on process management, and the horizontal constraint relationship is consistent. Therefore, the effect of standardization on process management can be extended to express the effect of standard system construction on process management (Figure 2).

First, the role of standardization. When standardization makes the space and time of the process management in a unified state, this unified state makes the process management in a time-invariant relationship or/and a space-invariant relationship, thereby forming a standardized role in process management or/and a standardized role in space . Timely unification of static process procedures and dynamic process order guarantees the standardized value of process recurrence. The spatial unification of the process constitutes the mutual standardization of different positions in the process. Standardization includes order, and the construction of a recursive standard system plays a role in ordering process management. Ordering includes a spatial order arranged according to the size of the space and a temporal order arranged according to the time sequence. For example, enterprise business processes are staggered, and at the same time, each business process cycle involves the joint participation of more departments, formulating interaction rules that are compatible with the business process and global integration capabilities, which regulates the flow relationship of the process; for complex processes, using standardized modularization concepts to decompose and extract functional programs, unify them into reusable process modules, and then unify the internal logic, status, and functions of the process modules, as well as the external characteristics of the corresponding interfaces, and so on. Standardize the quality behavior of the process module.

Second, the protection function. The safeguard function is a protection and defense function. After the implementation of standardization of process management, the unified state produced will make the specific effects of process management not change with the operator, work time, and location. The maintenance of this state plays a role in protection and defense for the input and output of process management. Therefore, the construction of the standard system has a safeguarding effect. For example: unifying the manufacturing process of products so that the manufacturing process of each product conforms to the unified eligibility requirements can guarantee the quality of each product output; unifying the process of implementing occupational hygiene control can help Different working hours and locations provide equivalent guarantees for the occupational health of the workers.

Third, saving effect. The saving effect of standardization comes from the saving effect brought by the unification of the process. This saving effect has economic saving effect and time saving effect respectively. The economic saving effect is the function of saving process management costs, and the time saving effect is the function of shortening the process time period.

For example, in the relationship between product development and production, the standard system is constructed to unify the specific attributes of similar products, so that products produced according to the same process have a wide range of generalization, thereby increasing the scope and quantity of the same products, and the number of products required. The increase in the production process has led to an increase in the number of applications of the production process, which in turn shared the process design cost and other sunk costs. Therefore, the construction of the standard system has played a cost saving in process management in terms of economic effects.

The saving effect of standardization has also played a driving effect on the efficiency increase of process efficiency. Efficiency is dependent on time, that is, the workload per unit time. The construction of the standard system fixes the best action of production labor. Repetitive labor will drive the improvement of proficiency, shorten the labor time under the same workload and

improve labor efficiency. Similarly, the increase of process application times in the same time will inevitably improve process application proficiency and process output efficiency. Therefore, the construction of standard system has shortened the process time cycle in terms of time effect.

3.2. The influence of process management on the construction of standard system

Process management pursues strict sequence limitation and statistical control from the input end to the output end. It is a continuous activity with a clear starting point and end point to achieve the optimal allocation of scarce resources. The standard system is the top-level design of the standard development, and the expression of the standard relationship system. The construction process of the standard system itself needs to reflect the systematic nature of the standard composition, matching coordination, forward-looking positioning, and advanced content from the input end. Therefore, process management can not only guide the construction of the standard system, but also check the purity of the state of the standard system. The following three points explain the driving effect of process management in promoting the construction of standard systems to realize value:

First, optimize the management of the standard system. Primarily, solve the problem of the integrity and uniformity of the standard. Through process management, all the activities of the production and operation management of the enterprise are determined, the activities and requirements of each node in the entire process are clarified, and the problems of standard-setting departmentalism and multiple issues of the same thing are solved, and it can ensure that the standard is in operation. The division of labor is implemented; secondly, the problem of poor standard coordination is solved. Through process management, the main responsible departments, participating departments, and implementation departments of each node of the process are clarified, and the responsibilities are stipulated in a standard form, which can effectively solve the problem of out of control and control of the joint. Finally, solve the shortcomings of standard compliance and operability. Through process management, all activities from the beginning to the end of each work and business process are defined in the form of flow chart. At the same time, the main responsible departments and personnel, participating departments and personnel, implementation departments and personnel of each node activity, as well as the documents to be executed and the records generated are specified, which can not only meet the requirements of superiors and actual work, And it is easy for the executor to implement.

Second, solve the "systematic degradation" of the standard system. Any standard system is designed based on the development environment, model and technical level at the time. After a period of development, it will become more and more exposed that the structure is not advanced and incomplete, and the standard age is too long, which leads to inapplicability and imprecise standard formulation work. This leads to systemic problems such as duplication, overlap, and contradiction among standards, and process management can reconstruct and optimize them. First of all, process management emphasizes timeliness and conveys a sense of crisis. Process management will monitor the output characteristic deviations in real time during the implementation of the standard. The external influences of the company, such as changes in regulations and policies, and the application of advanced technologies, will obviously lead to a mismatch in the process output interface. Process managers will find out and update the standard system in the first time. Secondly, the process management will review and reflect on the process regularly or irregularly, sort out, collect and rectify the problems encountered in the operation: whether the boundary of the process interface is clear, overlaps or cross-process simplification is feasible, the operability of the process, etc. The regulatory sub-cluster, risk sub-cluster, index sub-cluster, and post sub-cluster of the system are timely perfected and effectively combined.

Third, realize the flexible allocation of standard system construction resources. The fundamental motivation for the construction of the standard system is to meet people's subjective needs for the unique use nature of the standard. The subjective needs can be divided in order. At the same time, the construction resources of the standard system are scarce and constrained. It is biased to evenly allocate the resources to each module of the construction process of each standard system. Process management will give priority to the process modules that directly or indirectly contribute to the realization of the organization's strategic objectives under the coordination of the different, clear and implied needs and expectations of relevant stakeholders. At the same time, it will respond quickly according to environmental changes and prioritize according to the standard value demand relationship, so as to realize the flexible and reasonable allocation and supply of limited resources.

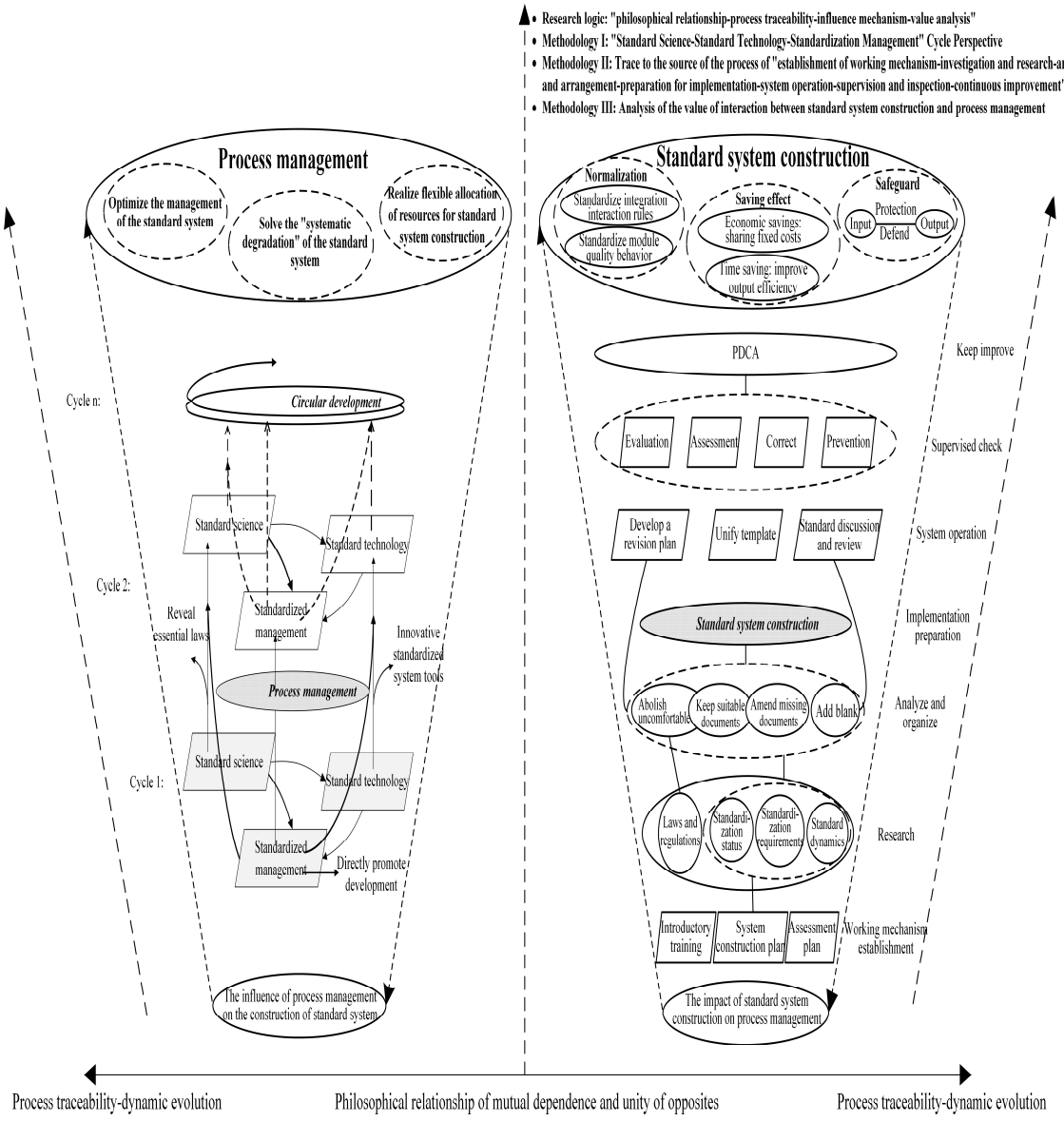

Figure 2. The mutual influence of standard system construction and process management

4. Model of Interaction and Coupling Relationship Between Standard System Construction and Process Management

4.1. Definition of coupling theory

Coupling theory is used to describe and measure the mutual affinity between two (or more than two) systems with similar properties, which promotes coordination, common development, and dynamic evolution of the systems.

The coupling relationship has three basic characteristics: self-organization, synergy, and measurability. Self-organization means that the coupled system can form a certain structure or function according to a certain rule among the various subsystems in the system without external instructions, so that the system as a whole changes from a chaotic and disordered initial state to a stable and orderly state. Final state evolution; synergy emphasizes the pulling effect of the cooperation between the coupled subsystems and coordination on the formation of the coupling as a whole, that is, in the evolution process of the mutual cooperation and harmonious development of the system, the cooperation of the subsystems makes the overall benefit index of the system "quantitative expansion" , subsystem coordination makes the overall structure of the system upgrade "quality improvement"; measurability refers to the orderly development process of the coupled system, the manager can use the data and information that can be obtained to describe the degree of phase change of the coupled structure and the order parameters. The strength and measurability of the synergy between each other enables verification and evaluation of the coupling level among the various subsystems within the system.

4.2. Standard system construction and process management interaction coupling relationship model composition

The expression of the coupling relationship between the construction of standard system and process management is different from the expression of the traditional relationship model, because the two are the organic sub-parts of the comprehensive management of the enterprise. The command and coordination of the "platform" of enterprise integrated management, so the expression of the coupling relationship between the two requires a comprehensive and comprehensive analysis of the coordinated changes between standard system construction and process management based on system theory, and the introduction of integrated enterprise management as the third provide the source of the coupling framework.

This paper divides the standard system construction and process management interaction coupling relationship model into a core system with a "people-regulation" composite system as the core, and a supporting platform provided by enterprise integrated management, standard system construction to standardize operating standards, and process management to stimulate endogenous motivation. The external system composed of the three, the components of the core, the external system, and the internal and external systems coordinate and respond to each other, and jointly drive and realize the interaction and coupling of standard system construction and process management (Figure 3).

4.2.1. Kernel system

The core system of the standard system construction and process management interaction coupling relationship model is a "human-regulation" composite system. Among them, "human" is the behavior subject of coordination and coupling, and "regulation" is the behavior object, which represents the management system to be optimized.

Acting subject-the participant and beneficiary of the standard system construction and process management. The behavior subject composed of the senior management team, project group, key experts and employees, is in the active position in the "people-regulation" relationship and runs through the transformation process. Motivation acts on "standardization" and "proceduralization", forming a direct driving force for coupling and interaction. Limited by the differences in willingness, resource shortage, and ability range, different actors play different roles in the coupling interaction model: the senior management team coordinates the direction of the coupling strategy through top-level design, path selection and resource allocation; As the basic work group of the enterprise, the project group uses its own cohesion and execution to guide the model coupling interaction and the realization of potential value through organization and coordination. The backbone elites often have a lot of management resources and can provide technology, intelligence and funds for the coupling interaction. Employees are the group most closely connected with the coupling interaction model. They have different requirements for the standard system structure and process maturity at different stages of enterprise development, and their decision-making willingness plays a decisive role in the coordination and interaction of the coupling model. In the process of coupling interaction, the degree of participation of different actors is different, and all parties jointly promote the construction of the standard system and the interaction of process management through the game of interests.

Behavior object-management system to be optimized. The object of behavior to be optimized is not suitable for the current corporate strategic needs or the current use of basic and reasonable institutional documents with potential disadvantages: On the one hand, there are problems such as inefficient use and passive management, and its original normative functions are gradually weakening or even declining. On the other hand, it has good corporate resources and value-added potential, but has not derived other functions through optimized configuration to meet the needs of corporate development.

The core system composed of "people" as the subject of behavior and "regulation" as the object of behavior is functioning well, realizing the dual-driven development of efficient teamwork and excellent systems within the enterprise.

4.2.2. Outer marginal system

The outer edge system is composed of many external factors that affect the operation of the core system. In view of the fact that the standard system construction and process management interaction coupling relationship model core system is guided by the integrated management of the enterprise, and there is an intersection of management behaviors between the integrated management of the enterprise and the process management, and the construction of the standard system, it should be placed in the coupling model. The top level of governance, process management, and standard system construction are placed at the grassroots level of policy management.

In the external system, the integrated management of the enterprise is the basic platform for the enterprise to maintain order. It first acts as a medium to realize the connection of resources for the market, upstream suppliers, and downstream distributors, and secondly promotes the quality characteristics of products or services and sustainable production. The organic coordination of the two benefits will finally realize the integration and development of advanced production technology and concepts, and guide the development direction of the internal source system of the coupled interactive model.

Process management is the subsystem that stimulates the endogenous power of the coupled system. It highlights the following three points: First, the redesign of process activities, which

realizes the reallocation and directional flow of limited resources, promotes the integration of enterprise processes and advanced technologies, and induces the core The behavior subject in the system acts on the driving force of the behavior object to achieve efficient management; the second is the optimization and standardization of process procedures to ensure that the same program is in different time relationships, different spatial relationships, and different operating objects, and can maintain the same presentation of idealized results; The third is the reshaping of the process system, which has caused many changes and transformations in the corporate quality management structure, environmental management structure, and occupational health protection structure, thereby inducing spontaneous changes in the coupled interaction model to a higher level.

As a subsystem of the coupling model, the standard system construction itself has standardization and standardization factors, and has long-term and stability. It is the norm and guarantee for the long-term stable operation of the coupled interactive model. At the same time, the standard system is constructed to build a bridge of mutual trust between the "customer" and the "business". The standard provides the supplier with the product quality requirements that the user expects, and the standard also provides the user with the quality that the supplier provides requirements, provide a basis for inspection and acceptance for both supply and demand parties, provide a basis for supervision of product market management, and realize the integration of mutual trust resources; secondly, the results of the standard system construction are implemented in the standard system table of the text, which is to understand and reproduce the production function of the enterprise The basic roots, and at the same time, it has learned the lessons from the previous enterprise operation and management process, so it is an explicit expression of the precipitation of enterprise experience; in addition, the construction of the standard system has always maintained the basic orientation of continuous optimization and dynamic development, ensuring the coupling model at this stage stability, while promoting the coupling model to a more scientific and efficient stage of improvement.

At the same time, due to the openness of the system, each factor of the subsystem can directly exchange material flow, value flow and information flow with the core system, and indirectly control the interaction direction of the coupling model by influencing the willingness, resources and capabilities of the actors.

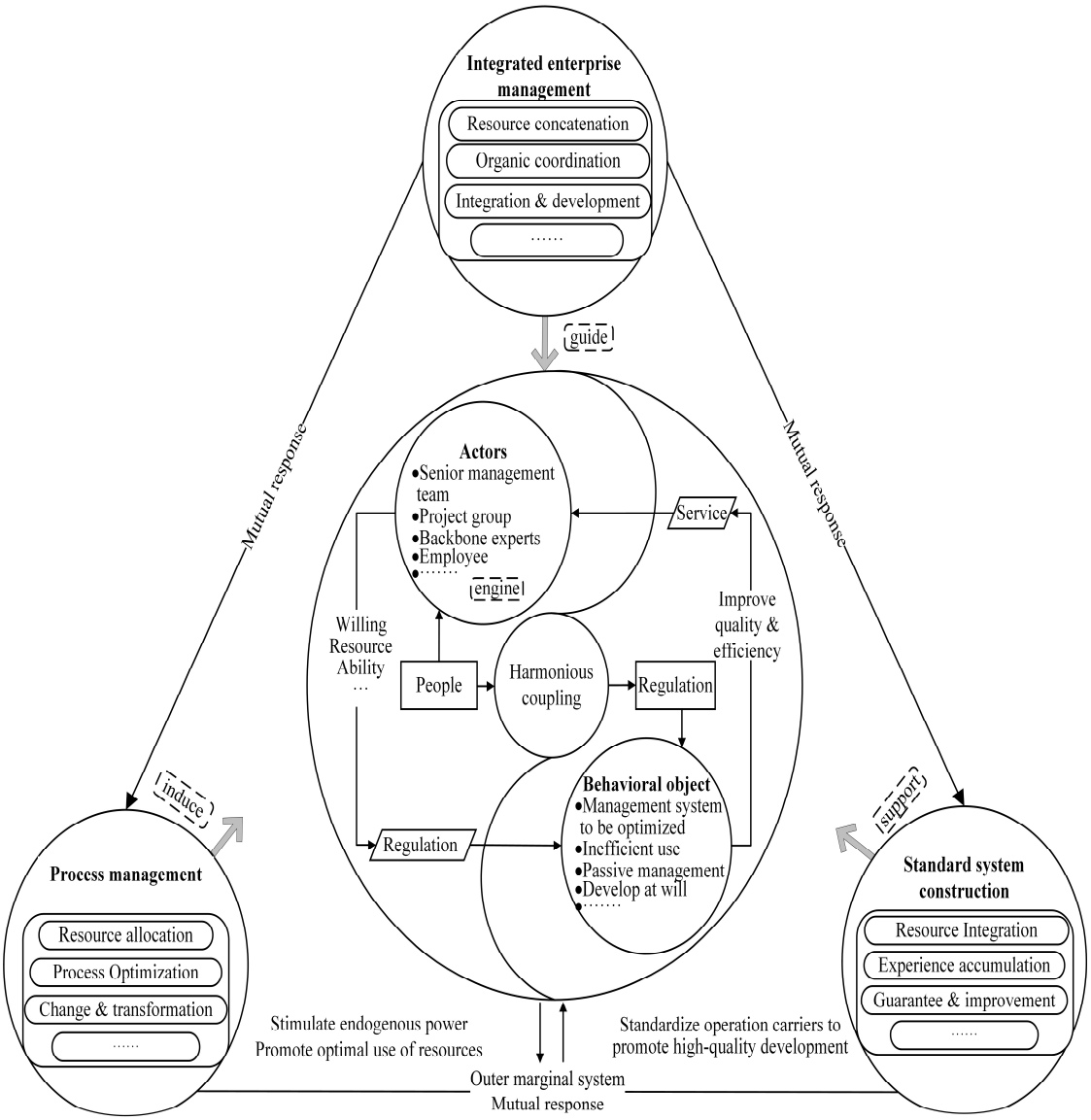

Figure 3. Model of interaction and coupling relationship between Standard system construction-Process management

5. THREE COUPLING MECHANISM CHARACTERISTICS

Analyzed from the perspective of management practice, the construction of standard system is the process of "realizing the unified state of the agreed scope", and the process management is the process of "guaranteeing the value reproduction of the forming sequence module". How to make the two coordinate in the coupling model It is more compatible and the coupling evolution process is smoother. The following is a list of three coupling mechanism characteristics that need attention:

5.1. The coupled system must be compatible with the dynamic deduction of the subsystem

Viewing the subsystem as a single system, the construction of the standard system itself is a complete standardization cycle, which is in a state of standardization and optimization at all

times; process management standardizes and systematically reorganizes and improves the process, while continuously improving the business performance of the organization , that is, it is in a state of process optimization. Both of them are in the process of their own dynamic deduction from this perspective, which also promotes the coupling model to spontaneously transition from a low-level stable stage to an unstable stage, and then develop into a higher-level stable stage.

5.2. The coupling effect has inertia

The coupling effect has the nature of state preservation, that is, the preservation ability of coupling "reform". The inertial nature of the coupling effect is embodied in two aspects. On the one hand, the coupling state is not easy to disappear; on the other hand, the coupling state is not easy to establish. The characteristics of the inertial nature of the coupling effect are: the inertial nature of the coupling effect is universal. As long as it is the state of the coupling effect, it must be a state of inertia; the magnitude of the coupling effect inertia is proportional to the coupling range and the time period of the coupling. The larger the coupling range and the longer the forming time period, the greater the inertia of the coupling state, and vice versa.

5.3. Use "system effect" to evaluate coupling benefits

First of all, the coupling benefit is the benefit produced by the application of the coupling concept. Coupling benefit is not the input and output relationship of the economic value of the coupling model, but the input and output relationship of the coupling concept. The benefit of the coupling model is not actually the contribution benefit of the model itself, but the benefit of applying the coupling concept and bringing about changes. The establishment of a coupling state requires all-round investment in personnel training, equipment, equipment, environment, and materials. From the perspective of knowledge, the coupling concept is a body of knowledge, not an economic reproductive body. It is not appropriate to use investment standards to evaluate the benefits of coupling; do not think too much about the best indicators of the sub-system, because the view of system science believes that the best sum of the sub-systems is not equal to the best total system, because the realization of the overall goal will bring "system effect", that is, the effect generated by the system as a whole exceeds the sum of the effects of various parts of the system.

REFERENCES

[1] Modern Chinese Dictionary, (2017) The Commercial Press, Dictionary Editing Office, Institute of Language Studies, Chinese Academy of Social Sciences,.
[2] ISO 9000 family, Quality management.
[3] Mai Lvbo, (2018) The method of standard system optimization. China Standardization, (07): 58-65.
[4] Wang Chaofan, Zhao Chaofei, (2021) Research on the construction of enterprise standard system from the perspective of methodology . China Standardization, (01): 90-94.
[5] GB/T 13016-2018, Principles and Requirements for the Construction of Standard System.
[6] Krislov, Samuel, (2020) How nations Choose Product Standards and Standards Change Nations, Pittsburgh: University of Pittsburgh Press.
[7] Harm Schepel, (2017) The Constitution of Private Governance: Product Standards in the Regulation of Integrating Markets, Oxford and Portland, Oregon.
[8] Jaesun Wang, Seoyong Kim, (2016) Time to get in: The contrasting stories about government interventions in information technology standards (the case of CDMA and IMT-2000 in Korea), Government Information Quarterly, 24.
[9] Smith G, (2015) Interaction of Public and Private Standards in the Food Chain, OECD Food, Agriculture and Fisheries Papers, No. 15, OECD publishing.
[10] Lawrence Busch, (2011) Standards: Recipes for Reality,Cambridge, MA and London: The MIT Press.

AUTHORS

Wang Chaofan is a postgraduate of management at North China Electric Power University, and currently serves as the general manager of Bohai Quality and Standardization Service Company.

RISK ANALYSIS IN THE PREPARATION OF A BUSINESS CONTINUITY PLAN (BCP) IN IT SERVICES: A CASE STUDY OF UNIVERSITAS INDONESIA

Akmal Gafar Putra, Betty Purwandari and Farisya Setiadi

Magister of Information Technology, Universitas Indonesia, Depok, Indonesia

ABSTRACT

Based on the Horizons Scan Report 2021 by BSI, the top 6 threats to organizations today are pandemics, health incidents, safety incidents, IT and telecommunications outages, cyber-attacks, and extreme weather. Universitas Indonesia (UI), as a modern, comprehensive, and open campus, strives to become a leading research university globally. As the IT service manager at UI, the Directorate of Information Systems and Technology (DSTI) has the task of strengthening service management by implementing risk management and security management in line with relevant laws and policies. The main problem for DSTI as an IT service at UI is that there are no documents related to risk management and information security management, resulting in IT services' failure. This year, there have been four data center failures due to power and UPS problems. DSTI wants to improve IT services at UI by implementing risk management and Business Continuity Management System (BCMS). This study aims to conduct a risk analysis to design a Business Continuity Plan (BCP) for IT services at the University of Indonesia. The research was conducted using mix method. The OCTAVE qualitative method was carried out in finding a list of risks on critical assets in IT services at UI. A quantitative approach is needed to rank the risk list using a questionnaire and FMEA calculations to get a risk priority number. This study separates the risk of general assets and information system assets. For critical assets, it is generally found that two are at a very high level, one is high, eight risks are at a low level, and 12 are at a very high level, for information system assets found 12 assets with very high risk, three medium and one low.

KEYWORDS

Risk Analysis, OCTAVE, FMEA, ISO 22301:2019, Business Continuity Plan.

1. INTRODUCTION

Universitas Indonesia, abbreviated as UI, is a modern, comprehensive, open, multi-cultural, and humanist university. UI simultaneously and continuously trying to become the world's leading research university. The vision stated in the statutes [1] is: "To become a center of science, technology, and culture" superior and competitive, through efforts to educate the nation's life to improve the welfare of the community, thereby contributing to Indonesian development people and the world. UI has developed three strategic goals that are expected to realize the goals of UI 2024. The three strategic targets in internal business processes are relevant and high-quality education, research-based Tridharma, and effective governance [2]. The Directorate of Information Technology and Systems, abbreviated as DSTI, is a directorate entrusted with being a trusted institution in the management of information technology infrastructure and data

David C. Wyld et al. (Eds): CSTY, SIGI, MaVaS - 2021
pp. 101-111, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.112209

processing to support the implementation of the Tridharma of Higher Education and the achievement of the UI development trilogy.

DSTI has a vision "To become an information system management institution with excellent service quality to support the achievement of a world-class UI, which is comparable to the management of information systems at other leading universities in Asia." DSTI has three missions: fostering reliable, integrated, secure, and information-rich information technology infrastructure from, by, and for the entire University of Indonesia academic community; realizing mature IT Governance for the University of Indonesia towards a world-class university; building and ensuring the realization of IT-based application system services and infrastructure that can assist management at UI by UI's strategic plan [2].

Based on a survey conducted by the British Standard Institution [3], the COVID-19 pandemic is the highest disruption compared to other incidents that occur in organizations; this is due to the organization's lack of preparation in dealing with this pandemic. Health incidents, another category considered low risk for 2020, ended the year as the second biggest nuisance. Many of these health incidents are not from pathological causes but from mental health difficulties experienced by staff due to COVID 19. Cyberattacks and IT/telecommunications outages are also causing high levels of disruption in 2020 due to increasing cybercrime. Criminals seek to exploit security loopholes while staff works remotely, and unexpected network outages are caused primarily by issues with internet latency. Based on the list, the first and second ranks are threats caused by COVID-19, followed by work safety incidents, IT and telecommunications blackouts, cyber attacks, and extreme weather in the top 6 positions.

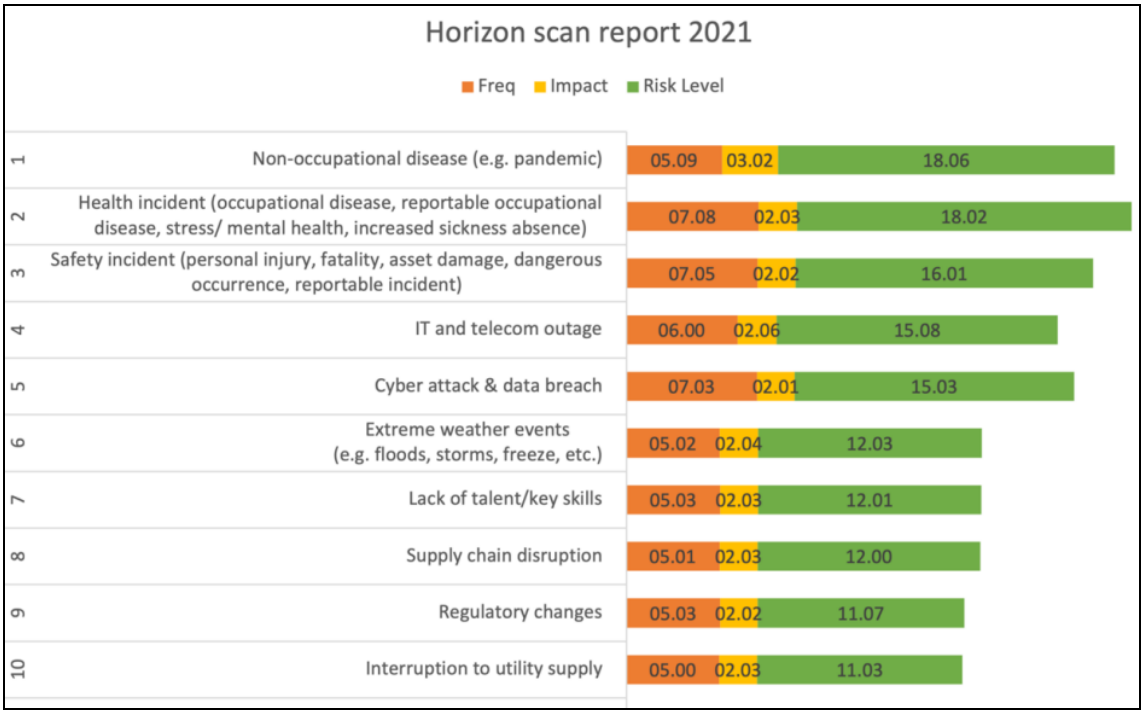

Figure 1. List of Threats to Organizations [3]

In a study conducted by Vasquez and Ortega, the loss of business-critical data and system downtime are two of the most significant risks faced by those in charge of information technology in the IT department [4]. And in other studies, it was found that the design of the Business Continuity Plan (BCP) serves to maintain the continuity of the company's business so

that it continues to run when the company's information technology experiences disruption,, where the research uses an international standard organization framework ISO 22301 [5].

Based on the results of interviews with the Information System Development Manager, there was a problem with the UI data center service, namely the server shutting down caused by damage to the Power Supply Unit and power generator failure for two days. The effect caused by the power failure was the data damage in the Storage Access Network which resulted in the blackout of several information systems for seven days.

From the analysis of the fishbone diagram, five domains affect the continuity of IT services at the University of Indonesia data center, namely: Currently, DSTI does not have a business continuity plan document, DSTI has two SysAdmins and two NetAdmin with levels competencies are in different level so that for disaster recovery DSTI is very dependent on individual staff, In pandemic conditions SysAdmin and NetAdmin often don't be at the data center location when a disaster occurs, so that data center recovery be disturbed, no backup or swap implementation yet to the main applications such as the academic system, financial system, staffing, and others to anticipate failures in the primary system or data center, The loss of the power source and the unavailability of a disaster recovery center are some infrastructure problems, company policy of 25% to 50% presence during a pandemic has a significant impact on disaster management.

This study aims to conduct a risk analysis to prepare a Business Continuity Plan (BCP) on IT services at the University of Indonesia. This research hopes that a list of risks needed to prepare the BCP will be obtained with this research.

2. LITERATURE REVIEW

2.1. Risk

Paul Hopkin defines risk as the possibility of harm, loss, or an accident that may occur [6]. Darril Gibson defines risk as a possible loss of a vulnerability that can threaten an organization [7].

2.2. Risk Management

Risk Management according to Darril Gibson is the practice of identifying, assessing, controlling, and reducing risk [7].

2.3. Business Continuity Management System

Business Continuity Management System or BMCS is the whole process managerial skills that can predict threats and their impacts on the organization if the hazard occurs and provide a framework and blueprint complete set of tools to build organizational resilience with relevant responses to protect key stakeholders, reputation, brand, and business process activities [8]. BCMS emphasizes the importance of understanding the organization's need to conclude policies and objectives, implement controls and estimates for the management of disturbing incidents, observe performance, and carry out sustainability efforts.

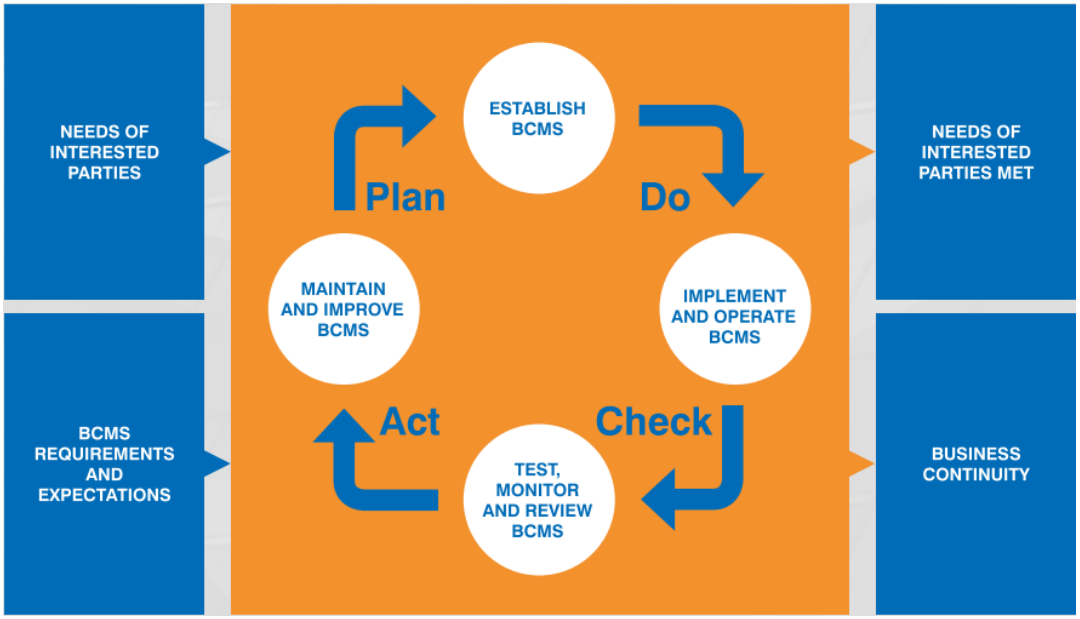

Figure 2. Plan, Do, Check, Act cycle in BCMS process

2.4. Business Continuity Planning

BCP is a document that contains instructions or procedures for how the organization ensures business processes during and after interruptions [9]. BCP is defined as a procedure document of organizational implementation guidelines to restore, repair, and restart basic operations in a state of disturbance [8]. The latest, according to Darril Gibson, BCP is a document used to help a company plan to disaster or emergency. The goal is to ensure that the operation critical organization continues to function. BCP consist of procedures and instructions used to restore operations in the event of a disaster.

2.5. Previous research

2.5.1. Business Continuity Plan in IT Solution Company (PT. ABC) Using ISO 22301:2012 [10]

This study aims to design BCP for companies engaged in solutions IT. This company has two data centers and a DRC. This study conducts a Risk Analysis in designing BCP using a framework work ISO.

2.5.2. Design of a business contingency plan. Case study: Municipality of Cantón Suscal. [4]

This study aims to communicate the results of the BCP design in the technological innovation (DIT) for the City of Cantón Suscal (MCS). This research analyses IT areas where the loss of business-critical data and system downtime are two risks biggest challenges faced by those in charge of information technology in the department of technological innovation (DIT) for the City of Cantón Suscal (MCS) (Vasquez & Ortega, 2020).

2.5.3. Business Continuity Plan Design and Technology Disaster Recovery Plan and Information Systems Using ISO 22301 [5]

This study conducts a risk analysis in the design of BCP with the framework ISO 22301 combined with BS 25999-1 and BS 25999-2.

2.5.4. Risk analysis on the development of a business continuity plan [11]

This research is about designing a business continuity plan that functions to maintain the continuity of the company's business so that it continues to run when technology information on the company is disrupted. This study uses international standard organization framework 22301 (I. Setiawan et al., 2019).

2.6. Risk Analysis Method

2.6.1. OCTAVE Method

OCTAVE (Operationally, Critical Threat, Asset and Vulnerability Evaluation) is a risk-based strategic assessment and planning technique for effective security developed at the Software Engineering Institute (SEI) [12]. OCTAVE is a framework that organizations can use to identify information security risks and help organizations develop risks qualitatively and identify critical assets to the organization's mission. OCTAVE targeted organizational risk and focused on strategic issues related to University of Indonesia practice. OCTAVE is a flexible evaluation that can be adapted to most organizations.

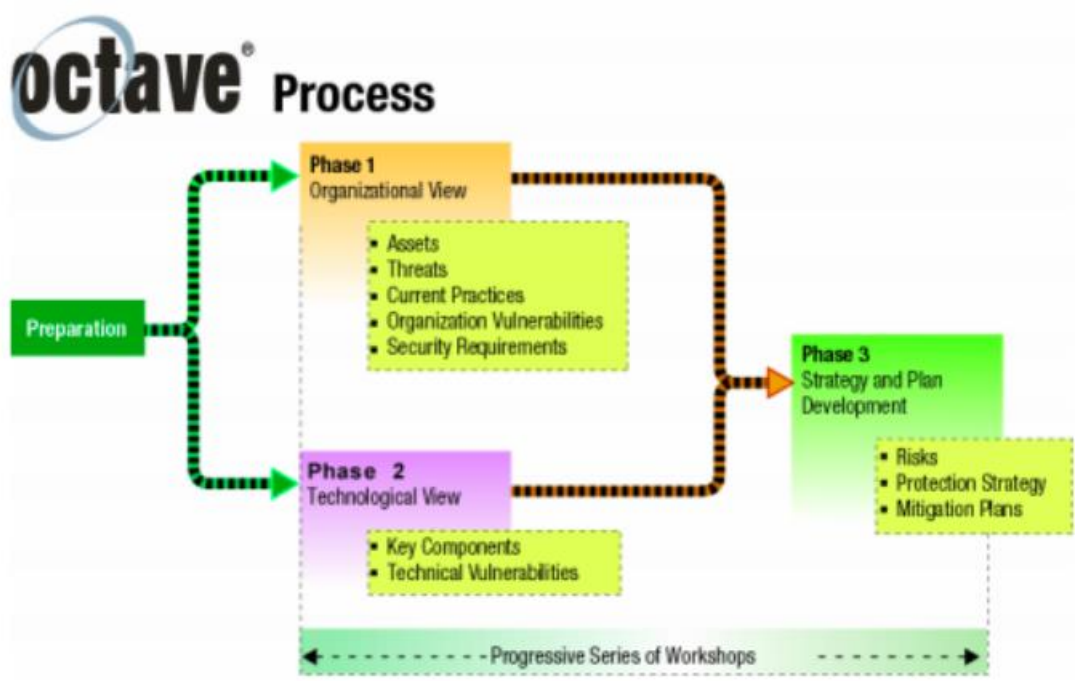

Figure 3. OCTAVE Method Process Flow

2.6.2. FMEA Method

FMEA (Failure Modes and Effects Analysis) is a systematic method used to identify the consequences or consequences of potential failures systems or process and reduce the chance of

failure [13]. FMEA is one of the reliable tools to minimize losses that occurred as a result of this failure. The 10 steps of the FMEA are; Identify the components and their related functions, Identify failure modes, Identify the impact of the failure mode, determine the severity of the failure, Identify the cause of the failure, Determine the frequency value of the occurrence of failure, Identify the necessary controls, Determine the effectiveness of the current control (detection), Calculate the RPN (risk priority number) value, and Determine actions to reduce failure.

In order to produce accurate output of risk analysis using the FMEA method, Firstly there must be some determination of the value of severity, occurrence, and detection. In the FMEA method, the final risk priority number is called the RPN or risk priority number. RPN is a mathematical result of the impact (Severity), the probability of occurrence of the risk that will lead to failure (Occurrence), and the ability to detect failures (Detection). The following equation shows the RPN value:

$$RPN = Severity \times Occurrence \times Detection$$

Table 1. Risk Level

Risk Level	**RPN Value**
Very High	>200
High	<200
Medium	<120
Low	<80
Very Low	<20

2.7. Comparison of Risk Analysis Methods

In this section, the differences in risk analysis methods between OCTAVE and NIST SP 800-30, the OCTAVE method of evaluating organizations, focuses on security practices within the organization, focusing on the organization's strategic issues and Self directions. the following table shows the difference in methods:

Table 2. Comparison of Risk Analysis Method

Method	**OCTAVE**	**NIST SP800-34 Rev.1**
Evaluation Subject	Organization	Information System
Focus	Security Practices	Technology Used
Issue	Organization Strategic	Organization Tactic
Direction	Self-Direction	Expert-Direction

3. ANALYSIS

This section discusses the stages of performing a risk analysis based on the method OCTAVE and FMEA

3.1. Asset Identification

At this stage, asset information is collected, grouped into 4 categories: Hardware, Software, Information, and Personal.

Table 3. Assets

Asset Code	Criteria	Asset Name
H1	Hardware	Server
H2		Firewall
H3		Laptop
H4		Monitor
H5		PC Unit
H6		Router
H7		Switch
H8		CCTV
H9		Electric Generator
H10		UPS
H11		Precise Air Conditioner
H12		Fire Distinguisher
H13		Cabel
I1	Information	Student
I2		Staff
I3		Finance
I4		Vendor
I5		Assets
S1	Software	OS
S2		LDAP
S3		SSO
S4		Ms. Office
S5		Database Tools
S6		App Dev Tools
S7		Academic
S8		Humar Resource
S9		Finance
S10		Logistic and Procurement
S11		Asset and Facilities
S12		Research
P1	Pegawai	Director
P2		Development Manager
P3		Operational Manager
P4		IT Gov Specialist

Asset Code	Criteria	Asset Name
P5		Solution Design Specialist
P6		App Dev Specialist
P7		App Operation Specialist
P8		Infra Dev Specialist
P9		Infra Operation Specialist
P10		Administration

3.2. Risk Assessment

Table 4. General Asset Risk Assessment

Category	**Critical Asset Name**	**Threat**	**S**	**O**	**D**	**RPN**	**Risk Level**	**Risk ID**
Hardware	Server	Hardware Failure	8	2	3	48	Low	1
		Run Out Resources	8	1	2	16	Very Low	2
		configuration crash	8	1	6	48	Low	3
	PC/ Laptop	Hardware Failure	5	1	1	5	Very Low	4
		Virus	5	1	1	5	Very Low	5
		Software Damage	5	1	1	5	Very Low	6
	UPS	Hardware Failure	8	1	1	8	Very Low	7
		Loosing Power	8	6	1	48	Low	8
	Electric Generator	Hardware Failure	8	1	6	48	Low	9
		Run Out of Gas	8	1	1	8	Very Low	10
Networking	Router & Switch	Hardware Failure	5	1	10	50	Low	11
		Configuration crash	5	1	10	50	Low	12
	Cable	Physical Damage	6	1	10	60	Low	13
	Firewall	Hardware Failure	2	1	10	20	Low	14
		Configuration crash	5	1	1	5	Very Low	15
Software	LDAP	Power Down	8	6	5	240	Very High	16
	SSO	Power Down	8	6	5	240	Very High	17
Data	Student	Inconsistencies Data	1	1	1	1	Very Low	18
	Human Resource	Inconsistencies Data	1	1	1	1	Very Low	19
	Finance	Inconsistencies Data	1	1	1	1	Very Low	20
	Research	Inconsistencies Data	1	1	1	1	Very Low	21

Category	Critical Asset Name	Threat	S	O	D	RPN	Risk Level	Risk ID
Human Resource	Staff	Resign	5	1	1	5	Very Low	22
		Work From Home	8	9	2	144	High	23

Table 5. Information System Risk Assessment

No.	Information System Name	Business Categories	S	O	D	RPN	Level RPN
1	emas2.ui.ac.id	Main	9	7	5	315	Very High
2	idols.ui.ac.id	Main	9	6	5	270	Very High
3	rima.ui.ac.id	Main	5	8	6	240	Very High
4	scele.ui.ac.id	Main	7	6	5	210	Very High
5	pra-registrasi.ui.ac.id	Main	5	4	10	200	Very High
6	bp.ui.ac.id	Main	6	3	10	180	High
7	edom.ui.ac.id	Main	7	5	5	175	High
8	evisem.ui.ac.id	Main	7	5	5	175	High
9	emas.ui.ac.id	Main	9	6	3	162	High
10	academic.ui.ac.id	Main	7	2	10	140	High
11	beasiswa.ui.ac.id	Main	5	5	5	125	High
12	pdf.midearth	Main	9	1	10	90	Medium
13	pjj.ui.ac.id	Main	6	5	3	90	Medium
14	ovis.ui.ac.id	Main	6	5	2	60	Low
15	Lontar (Perpustakaan)	Support	7	6	8	336	Very High
16	Remote-Lib	Support	7	6	8	336	Very High
17	Unggah.ui.ac.id	Support	7	6	8	336	Very High
18	Assets and Facilities	Support	5	6	10	300	Very High
19	Mailing	Support	7	7	6	294	Very High
20	Redmine	Support	7	5	8	280	Very High
21	Finance	Support	7	6	5	210	Very High
24	Human Resource	Support	5	5	5	125	High
25	HRIS	Support	5	5	5	125	High
26	UI Archive	Support	5	4	6	120	High
27	Vehicle Loan	Support	5	6	4	120	High
28	Arsip	Support	5	4	5	100	Medium

4. CONCLUSIONS

This study aims to conduct a risk analysis in designing the BCP Business Continuity Plan to be implemented in IT services at the University of Indonesia. The conclusion obtained is as following: At the beginning of 2020, the threat of COVID-19 emerged as a global threat that affect the list of threats and risks to business services in IT services; DSTI UI has not conducted a comprehensive risk assessment related to hardware, software, information, and employees asset; OCTAVE method is suitable for identifying all critical assets, critical asset components, threats, vulnerabilities to key components, practices security, and organizational vulnerabilities; The FMEA method is suitable for measuring the value of risk if it already exists standard measurement methods for the impact, incidence and detection of each risk profile; Found several threats of power source failure such as UPS and Genset have a high and systemic impact on other assets. On system assets information there are 12 SI with very high risk level, 12 high risk, 3 medium and 1 low; the final conclusion is that the risk analysis process in this study has produced a list of risks needed in the preparation of the BCP which contains IT management policies in disaster preparation and management.

ACKNOWLEDGEMENTS

The authors would like to thank everyone, just everyone!

REFERENCES

[1] Statuta UI. Statuta Universitas Indonesia, Pub. L. No. 68 (2013). Indonesia.
[2] Renstra UI. (2017). Rencana Strategis Universitas Indonesia Revisi 2015-2019, 1–90.
[3] Business Continuity Insitute. (2021). BCI Horizon Scan Report 2021. 40, 55.
[4] Vasquez, E. J., & Ortega, J. C. (2020). Design of a business contingency plan. Case study: Municipality of Cantón Suscal. 2020 International Conference on Intelligent Systems and Computer Vision, ISCV 2020. https://doi.org/10.1109/ISCV49265.2020.9204334
[5] Setiawan, I., Waluyo, R., & Pambudi, W. A. (2019). Perancangan Business Continuity Plan dan Disaster Recovery Plan Teknologi dan Sistem Informasi Menggunakan ISO 22301. Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), 3(2), 148–155
[6] Hopkin, P. (2017). Fundamental of Risk Management Understanding, evaluating and implementing effective risk management. (4th, Ed.). KoganPage.
[7] Gibson, D. (2014). Managing Risk in Information Systems. Retrieved from https://books.google.com/books?id=5lktBAAAQBAJ&pgis=1
[8] ISO 22301 (2012). Social Security – Business Continuity Management Systems Requirement. ISO.
[9] NIST. (2010). NIST Special Publication 800-34 Rev.I: Contingency Planning Guide for FederalIinformation System. Gaithersburg: NIST..
[10] Pramudya, G. W., & Fajar, A. N. (2019). Business continuity plan using ISO 22301:2012 in IT solution company (pt. ABC). International Journal of Mechanical Engineering and Technology, 10(2), 865–872.
[11] Setiawan, A., Wibowo, A., & Susilo, A. H. (2018). Risk analysis on the development of a business continuity plan. Proceedings of the 2017 4th International Conference on Computer Applications and Information Processing Technology, CAIPT 2017, 2018-Janua, 1–4. https://doi.org/10.1109/CAIPT.2017.8320736
[12] Alberts, C., Dorofee, A., Stevens, J., & Carol Woody. (2003). Introduction to the OCTAVE® Approach. August, 121–129. https://doi.org/10.1016/b978-0-7020-3055-0.00004-2
[13] Stamatis, D. H. (1996). Failure Mode and Effect Analysis: FMEA From Theory to Execution. In Technometrics (Vol. 38, Issue 1). https://doi.org/10.1080/00401706.1996.10484424

AUTHORS

I am a master program student at the Faculty of Computer Science, University of Indonesia.

AUTHOR INDEX