Selected Computing Research Papers

Volume 9

June 2020

Dr. S. Kendal (editor)

Published by
the
University of Sunderland

Edited, typeset and printed by
Dr. S Kendal
University of Sunderland
David Goldman Informatics Centre
St Peters Campus
Sunderland
SR6 0DD

Tel: +44 191 515 2756
Fax: +44 191 515 2781

Contents Page

Critical Evaluation of Vocal Disorder Detection Methods (Nino Bahn).............................. 1

Critical Evaluation of object recognition models and Their Speed and Accuracy on Autonomous Vehicles (Luis Castrillo) 9

A Critical Evaluation of Artificial Intelligence Research Aimed At Improving Robotic System To Help Visually Impaired People (Oussama El Jerche).......................... 17

Critical Evaluation of Current Methods Aimed at Improving Ransomware Detection (Thomas Herb)... 23

An Evaluation of Defence Methods against HTTP DoS Attacks (Hieu Nguyen).............. 31

A Critical Evaluation of Current Techniques Aimed at Improving Physical Layer Security in Wireless Network (Lip Chuan Sui) 37

Critical Analysis of Current Research Aimed at Improving Detection of Phishing Attacks (Eddie Wong Shou Shie) 45

Critical Evaluation of Vocal Disorder Detection Methods

Nino Bahn

Abstract

With the ever-evolving technology of artificial intelligent, non-invasive methods in tracking pathological voices has become to be an important subject in latest research. This paper focuses on Vocal Disorder Detection Methods, by analyzing different classification approaches through machine learning techniques from current research papers. Finally, comparisons and conclusions are reached from the discussed methods and test results with consideration of their reliability and methodology. Also, a suggestion for future work with a combination of several methods will be presented, for an efficient and versatile mobile system to detect various diseases.

1 Introduction

A Vocal disorder is a significant health issue and wildly spread among many people. It effects especially professionals like singers or teachers who need to rely on their voice. "Untreated, voice disorders cost billions of dollars in lost productivity. For an occupation like teaching where voice use is heavy, the cost is almost $3 billion annually." (Ilapakurti et.al. 2019). Therefore, a non-invasive method which is reliable and accurate in detecting vocal disorders through a mobile device would help people to save time and money, having the possibility of an early diagnosis without consulting a professional.

Shia and Jayasree (2017) used Discrete Wavelet Transform and Feed Forward Neural Network in order to build a voice disorder detection system and showed, that energy of wavelet subband coefficients can be used as a reliable feature for detection of voice pathology.

Also, Sripriya et.al (2017) proposed a research on detecting pathological voices through analysing different types of jitter and shimmer measurements and using the Dynamic Programming Phase Slope algorithm.

Lopes et.al. (2018) analysed the accuracy of Recurrence Qualification Measurements in their study in order to discriminate voices with and without disorders.

This research paper will analyse the methods and experiments from current research in detecting voice disorders using different machine learning approaches for classification through critical evaluation in order to find the most effective methods with consideration of the detection accuracy and methodology. Also, possible comparisons and combinations of these methods will be discussed.

2 Current Voice Disorder Detection Techniques using different Classifiers

In this section current methods for Voice Disorder Detection using different classifiers from several research papers in particular Support Vector Machine, Boosted Tree, Deep Neural Network, Discriminative Paraconsistent Machine and K-Nearest Neighbours, will be analysed with a critical evaluation of their experiments and results.

2.1 Support Vector Machine

Amami and Smiti (2017) used a modified Density Based Spatial Clustering of Applications with Noise (DBSCAN) in order to detect noisy voice samples and an output model submitted to a Support Vector Machine (SVM) classifier with a radial basis function kernel to distinguish between normal and pathological voices.

For the experiment, Amami and Smiti (2017) used voice samples from 53 healthy speakers and 173 samples from speakers with pathological voices with a similar gender and age and different voice pathologies withdrawn from the Massachusetts Eye and Ear Infirmary Voice and Speech Laboratory database for the voice data. The voice samples are divided into groups grouped by a modified DBSCAN clustering method in order to detect the noises more efficiently with a dynamic ability to an overtime evolving incremental voices database with the ability of detection noises. Then, the SVM classifier determines the voices between pathological and non-pathological. Figure 1 illustrates the method. The results of the experiment showed a detection accuracy of 98%.

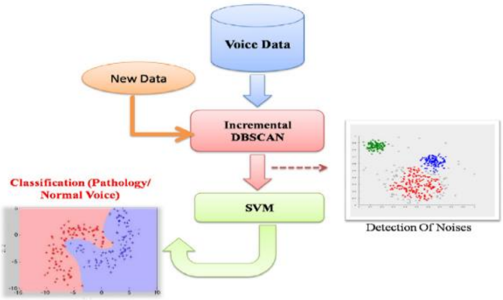


Figure 1 – Voice disorder detection using incremental DBSCAN and SVM (Amami and Smiti 2017)

The proposed approach from Amami and Smiti (2017) showed an efficient method in detecting vocal disorders by noise clustering and updating new changes to the training set without repeating the whole process.

In order to classify voices to normal and pathological, Guan and Lerch (2019) proposed four different approaches using deep learning and feature learning methods, these include (1) support vector machine (SVM), (2) Convolutional Neural Network (CNN), (3) CNN followed by SVM, and (4) autoencoder (AE) followed by SVM, as shown on Figure 2. Also, worth mentioning is the method of Alhussein and Muhammad (2018) with promising results using CNN with SVM in a mobile system to detect vocal disorders.

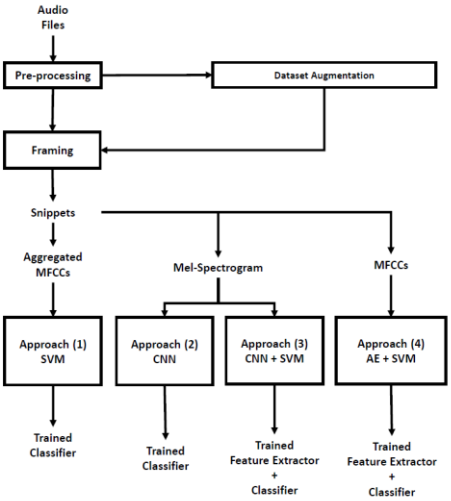


Figure 2 – Voice classification using deep learning and feature learning methods (Guan and Lerch 2019)

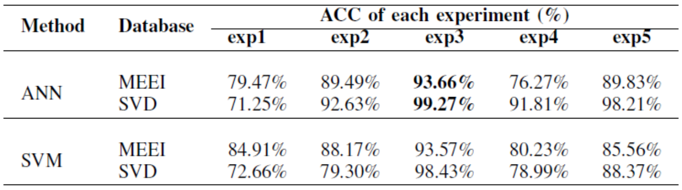
Guan and Lerch (2019) performed four experiments using voice samples from 2 datasets, the Massachusetts Eye and Ear Infirmary (MEEI) and the Technical University of Madrid (UPM) dataset. The experiments were divided into non-augmented and augmented data of each dataset. Also, a 5-fold cross-validation was applied on all experiments. The results showed acceptable improvements from experiments with non-augmented to augmented data on each dataset with the highest accuracy of 95.9% on the augmented MEEI dataset.

Guan and Lerch (2019) experiments illustrated valid science practice done in a controlled environment by using 2 independent datasets with identical conditions throughout every experiment, for evaluating their system.

Ezzine and Frikha (2018) used Artificial Neuron Network (ANN) and Support Vector Machines (SVM) in their research to classify vocal pathologies especially focused on benign and malignant tumours by investigating the performance of glottal flow features.

For the experiment, Ezzine and Frikha (2018) used the Massachusetts Eye and Ear Infirmary (MEEI) database and Saarbruecker Stimmdatenbank (SVD) database. From the MEEI database, voice samples from 52 normal and 657 pathological voices where being used. And 869 normal and 1356 pathological from the SVD dataset. Each dataset included a variety of different ages and genders including patients who have been diagnosed with benign and malignant tumours. Ezzine and Frikha (2018) performed five experiments, each using a various number of glottal features evaluated by using the classifiers ANN and SVM. Table 1 shows the accuracy result of each testing with ANN having the best results on experiment 3 with an accuracy of 93.66% on MEEI and 99.27% on SVD the dataset.

Ezzine and Frikha (2018) research showed good science practice with a valid approach on the experiments using glottal flow features in order to distinguish voice pathologies including benign and malignant tumours with results showing very high accuracy values.



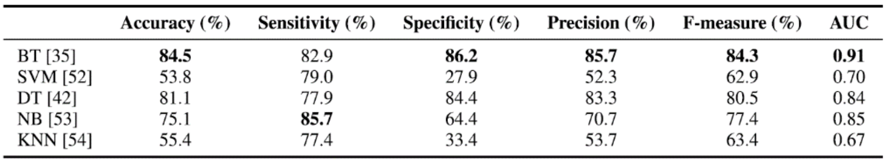
Method	Database	ACC of each experiment (%)				
		exp1	exp2	exp3	exp4	exp5
ANN	MEEI SVD	79.47% 71.25%	89.49% 92.63%	**93.66%** **99.27%**	76.27% 91.81%	89.83% 98.21%
SVM	MEEI SVD	84.91% 72.66%	88.17% 79.30%	93.57% 98.43%	80.23% 78.99%	85.56% 88.37%

Table 1 – Results in accuracy of the experiments using ANN and SVM (Ezzine and Frikha 2018)

2.2 Boosted Tree

Verde et.al. (2019) proposed a voice disorder detection system using the Boosted Tree (BT) algorithm as a classifier on a mobile device in order to handle the distinction between a healthy and a pathological voice.

The machine learning algorithm is trained on an external server and has been implemented in a mobile application. As shown on Figure 3, the howl system is divided into two main components, an external server which includes a BT algorithm used as and trained by voice samples from three different databases, and a personal mobile device. The trained algorithm was than implemented into the mobile device as a classifier.



	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F-measure (%)	AUC
BT [35]	**84.5**	82.9	**86.2**	**85.7**	**84.3**	**0.91**
SVM [52]	53.8	79.0	27.9	52.3	62.9	0.70
DT [42]	81.1	77.9	84.4	83.3	80.5	0.84
NB [53]	75.1	**85.7**	64.4	70.7	77.4	0.85
KNN [54]	55.4	77.4	33.4	53.7	63.4	0.67

Table 2 – Comparison of the results obtained with several classifiers (Verde et.al. 2019)

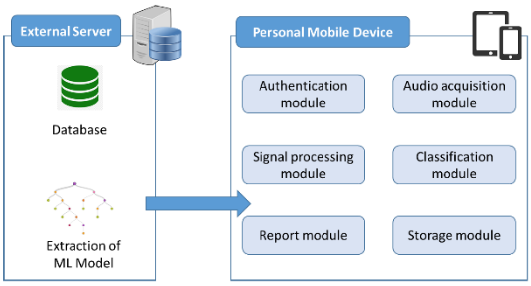


Figure 3 – Voice disorder detection system (Verde et.al. 2019)

In the experimental phase Verde et.al. (2019) carried out several tests using voice samples of 2003 voices extracted from three different databases; the Massachusetts Eye and Ear Infirmary (MEEI), the Saarbruecken Voice Database (SVD) and the VOice ICar fEDercio II (VOICED) database in order to test the reliability of the classifier.

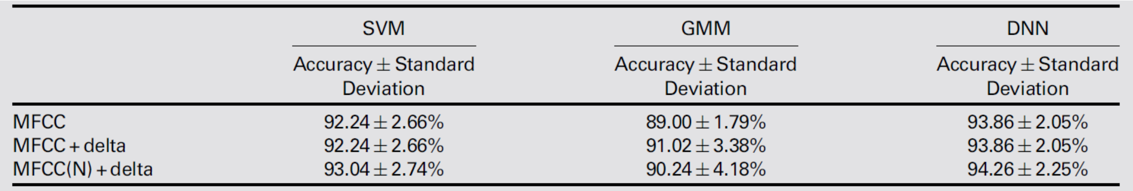
Verde et.al. (2019) also compared the classification performance of the BT algorithm with four other algorithms similar characteristics which are Support Vector Machine (SVM), single Decision Three (DT), Naive Bayes (NB) and K-nearest neighbour (k-NN). Table 2 illustrates a comparison of the results, with BT having the highest values in accuracy, specificity, precision, F-measure and AUC and only NB having the best value on sensitivity.

Verde et.al. (2019) concluded that their approach on using the BT algorithm as a classifier in order to detect vocal disorders has the best accuracy compared with other main machine learning algorithms for their proposed mobile voice disorder detection system.

The experiment is well justified in a controlled environment through the use of different and well-balanced datasets an. The performed tests indicate no bias while comparing the different algorithms with each other. Also, the proposed system shows a smart solution for managing the hardware performance as well as providing a certain level of privacy due to the fact, that no personal data is being sent to an external server.

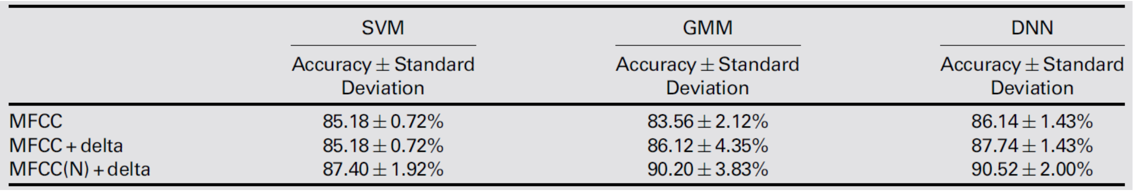
2.3 Deep Neural Network

In order to determine vocal disorders Fang et.al. (2019) used the Deep Neural Network (DNN) algorithm using three Mel frequency cepstral coefficient features (MFCC, MFCC + delta, MFCC(N) + delta) in comparison with two other algorithms, namely, Support Vector Machine (SVM) and Gaussian Mixture Model (GMM).
For the experiment, Fang et.al. (2019) obtained voice samples from the voice clinic in a tertiary teaching hospital (Far Eastern Memorial Hospital, FEMH). These samples include 60 normal voices and 402 voices of 8 common clinical voice disorders, divided into 205 male voices including 16 normal and 189 pathological voices, and 257 female voices including 44 normal and 213 pathological voices. Also, the Massachusetts Eye and Ear Infirmary (MEEI) database was used to verify the performance of their proposed method.



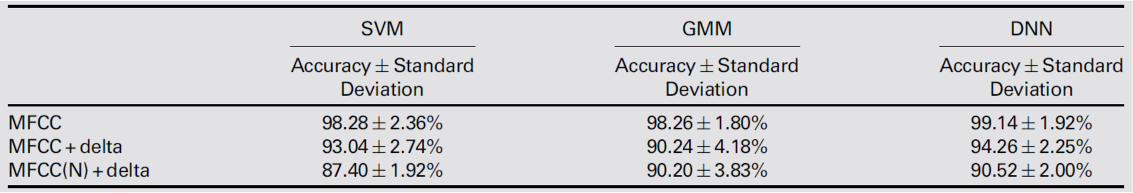
	SVM	GMM	DNN
	Accuracy ± Standard Deviation	Accuracy ± Standard Deviation	Accuracy ± Standard Deviation
MFCC	92.24 ± 2.66%	89.00 ± 1.79%	93.86 ± 2.05%
MFCC + delta	92.24 ± 2.66%	91.02 ± 3.38%	93.86 ± 2.05%
MFCC(N) + delta	93.04 ± 2.74%	90.24 ± 4.18%	94.26 ± 2.25%

Table 3 – Classification Accuracies of Three Classification Algorithms and Three MFCC Features Among Male Subjects (Fang et.al. 2019)



	SVM	GMM	DNN
	Accuracy ± Standard Deviation	Accuracy ± Standard Deviation	Accuracy ± Standard Deviation
MFCC	85.18 ± 0.72%	83.56 ± 2.12%	86.14 ± 1.43%
MFCC + delta	85.18 ± 0.72%	86.12 ± 4.35%	87.74 ± 1.43%
MFCC(N) + delta	87.40 ± 1.92%	90.20 ± 3.83%	90.52 ± 2.00%

Table 4 – Classification Accuracies of Three Classification Algorithms and Three MFCC Features Among Female Subjects (Fang et.al. 2019)



	SVM	GMM	DNN
	Accuracy ± Standard Deviation	Accuracy ± Standard Deviation	Accuracy ± Standard Deviation
MFCC	98.28 ± 2.36%	98.26 ± 1.80%	99.14 ± 1.92%
MFCC + delta	93.04 ± 2.74%	90.24 ± 4.18%	94.26 ± 2.25%
MFCC(N) + delta	87.40 ± 1.92%	90.20 ± 3.83%	90.52 ± 2.00%

Table 5 – Detection of Pathological Voice Samples in the MEEI Voice Disorder Database (Fang et.al. 2019)

As illustrated in Table 3, 4 and 5, the results of Fang et.al. (2019) experiment in detecting voice pathologies reached 94.26% of accuracy in male subjects, 90.52% in female subjects and 99.32% with the MEEI database using DNN as a classifier.

Fang et.al. (2019) concluded in their research, that using a DNN classifier and MFCC features outperforms a GMM and SVM classifier in terms of the accuracy. Also, as a future work suggestion, Fang et.al. (2019) proposed offline model (recording followed by analysis) can be developed into an online model (simultaneous recording and analysis) where voice samples could be analysed in real time using a cloud computing system.

Fang et.al. (2019) research showed well performed science in a controlled environment through justifying their proposed method in using DNN as a classifier in comparison with SVM and GMM using the same criteria with identical datasets.

2.4 Discriminative Paraconsistent Machine

Fonseca et.al. (2020) proposed a method to detect voice disorders by using the concepts of signal energy (SE), zero-crossing rates (ZCR) and signal entropy (SH), which provide a joint time-frequency-information map, in combination with the discriminative paraconsistent machine (DPM) for a voice signal-based classification.

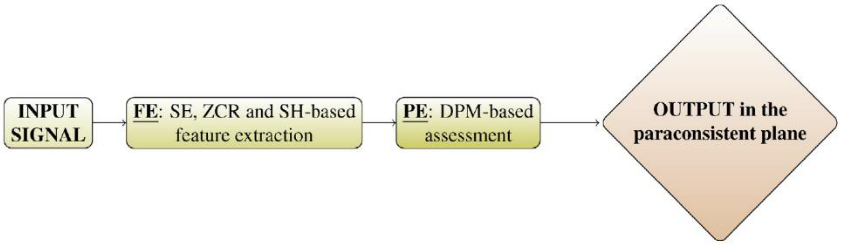

INPUT SIGNAL
FE: SE, ZCR and SH-based feature extraction
PE: DPM-based assessment
OUTPUT in the paraconsistent plane

Figure 4 – proposed method, where FE and PE are the active parts (Fonseca et.al. 2020)

The Fonseca et.al. (2020) approach was based on a feature extraction (FE) step containing SE, ZCR and SH in which the voice signals where getting fed into, followed by a DPM-based paraconsistent evaluation (PE) in order to treat indefinitions and contradictions, as illustrated in figure 4. For the testing the Saarbruecken voice database has been used with 687 healthy voice samples and 126 samples from patients which had Reinke edema, laryngitis and both Reinke edema and laryngitis.

Fonseca et.al. (2020) testing results concluded an accuracy value of 95% with handling indefinitions and contradictions where no pathologically-affected subjects where misclassified as being healthy. As a basic comparison, two Support Vector Machines (SVM) with radial basis function kernels where created in order to classify either Reinke edema or laryngitis which undertook the same procedure as DPM as the input. The SVM approach achieved a result of a 90% accuracy.

The proposed method of Fonseca et.al. (2020) based on a feature extraction including the combination of SE, ZCR and SH, associated with DPM, showed an efficient approach in vocal disorder detection with a high accuracy value in the testing. Also, Fonseca et.al. (2020) successfully validated their choice of DPM as a classifier with two SVMs which resulted with a lower accuracy in the testing than with DPM. Furthermore, Fonseca et.al. (2020) mentioned future research with the focus on experimenting DPM with more extensive datasets including voice samples with various disorders.

2.5 K-Nearest Neighbours

Mohamed and Guerti (2018) proposed a method on detecting and classifying neurological voice disorders based on parameters extracted from glottal signals using K-Nearest Neighbours (KNN) as a classifier.

In the proposed method, Mohamed and Guerti (2018) first used an Interactive Adaptive Inverse Filtering (IAIF) technique of the voice signal to estimate the glottal signal and time domain parameters used as features input vector to the KNN classifier as illustrated at figure 5. The classification of the results is divided into normal voice, Parkinson disease and spasmodic dysphonia.

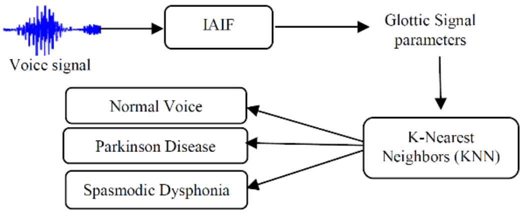


Figure 5 – Method of Neurological Voice Detection (Mohamed and Guerti 2018)

For a dataset to perform the experiment Mohamed and Guerti (2018) used the Saarbruecken Voice Database (SVD). From the database, 120 voice samples are selected of which 75% was used for the training phase and 25% for the testing phase. "In this configuration, the input vector of classifier comprising 12 parameters extracted from glottal signal is evaluated as input features set to KNN classifier." (Mohamed and Guerti 2018).

Voices	Normal	PD	SD
Training	30	30	30
test	10	10	10
Correct classification	10	9	9
Rate of classification	100%	90%	90%
Sensitivity	90 %		

Table 6 – Classification Results (Mohamed and Guerti 2018)

As shown in table 6, the method detected all normal voices correctly and 90% of accuracy each on Parkinson disease and spasmodic dysphonia.

Mohamed and Guerti (2018) conclusion states that KNN has a higher efficiency with the use of Glottal signal Parameters as input vector with an overall result of 93,33% accuracy during the experiment. Their research showed, that Glottal Signal Parameters features can not only be useful for detecting neurological voice disorders but can also be used to determine different types of vocal fold pathologies.

Mohamed and Guerti (2018) study overall indicate good science practice with one exception; more data from different datasets could have been used for a better justification in reliability of their proposed method. As an example, voice samples from the Massachusetts Eye and Ear Infirmary database or voice samples from a voice clinic with similar disorders from the experiment could have been collected and implemented in their testing.

3 Comparison of Methods

A comparison of the discussed methods should not only focus on the results in terms of accuracy but also in reliability, methodology and overall proposed system as the used algorithms differ in how they work.

In terms of reliability, the method of Ezzine and Frikha (2018) first seam to have promising results with a 99.27% prediction accuracy, but only covering the detection of benign and malignant tumours through the use of glottal flow features. The approach of Amami and Smiti (2017) on the other hand, got an accuracy of 98% covering a wider detection range of vocal pathologies with the second highest accuracy of every discussed method in this paper.

For practicality and patient privacy, the presented mobile system from Verde et.al. (2019) has the most practicable and affordable use in order to reach a wide range of patients with a diagnosis through a mobile device.

4 Conclusions

Every method discussed in this paper showed good science practice and promising approaches in order to detect various diseases through vocal signals with the individual experiment showing high accuracy results. Worth to mention is, that Fonseca et.al. (2020) and Mohamed and Guerti (2018) managed to detect every non-pathological subject correctly. Also, as Ezzine and Frikha (2018) were able to detect benign and malignant tumours with a particular high detection accuracy. Their method could be used to build an extended version of it, which can detect a wilder range of different or similar diseases.

For a future work suggestion, the proposed mobile system from Verde et.al. (2019) could be combined, with an upgraded version of the methods from Amami and Smiti (2017) for reliable voice disorder detection, Ezzine and Frikha (2018) for the detection of benign and malignant tumours and Mohamed and Guerti (2018) to classify Parkinson disease and spasmodic dysphonia. With that, a wild range of diseases could be detected with a high classification accuracy built in an efficient and affordable mobile system particularly for potential patients with consideration of their privacy.

References

Alhussein M, Muhammad G, 2018, ‘Voice Pathology Detection Using Deep Learning on Mobile Healthcare Framework’, *IEEE Access Volume 6*, Pages 41034–41041

Amami R, Smiti A, 2017, ’An incremental method combining density clustering and support vector machines for voice pathology detection’, *Computers and Electrical Engineering Volume 57*, Pages 257–265

Ezzine K, Frikha M, 2018, ‘Investigation of Glottal Flow Parameters for Voice Pathology Detection on SVD and MEEI Databases.’, *2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*

Fang S-H, Taso Y, Hsiao M-J, Chen J-Y, Lai Y-H, Lin F-C, Wang C-T, 2019, ‘Detection of Pathological Voice Using Cepstrum Vectors: A Deep Learning Approach’, *Journal of Voice, Vol. 33, No. 5,* Pages 634–641

Fonseca ES, Guido RC, Barbon Junior S, Dezani H, Gati RR, Pereira DCM, 2020, ‘Acoustic investigation of speech pathologies based on the discriminative paraconsistent machine (DPM)’, *Biomedical Signal Processing and Control*, 55, Pages 8-15

Guan H, Lerch A, 2019, ‘Learning Strategies for Voice Disorder Detection’, *2019 IEEE 13th International Conference on Semantic Computing (ICSC)*, Pages 295-301

Ilapakurti A, Kedari Sh, Vuppalapati JS, Kedari Sa, Vuppalapati C, 2019, ‘Artificial Intelligent (AI) Clinical Edge for Voice disorder Detection’, *2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService)*, Pages 340-345

Lopes LW, Vieira VJD, Costa SL do NC, Correia SÉN, Behlau M, 2018, ‘Effectiveness of Recurrence Quantification Measures in Discriminating Subjects With and Without Voice Disorders’, *Journal of Voice*, article in press

Mohamed D, Guerti M, 2018, ‘Glottal Signal Parameters as features set for Neurological voice disorders diagnosis using K-Nearest Neighbors (KNN)’, *2018 2nd International Conference on Natural Language and Speech Processing (ICNLSP) Natural Language and Speech Processing (ICNLSP)*, Pages 1-5

Shia S. Emerald, Jayasree T, 2017, ’Detection of pathological voices using discrete wavelet transform and artificial neural networks’, *2017 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS) Intelligent Techniques in Control, Optimization and Signal Processing (INCOS)*, Pages 1-6

Sripriya N, Poornima S, Shivaranjani R, Thangaraju Preethi, 2017, ‘Non-Intrusive Technique for Pathological Voice Classification using Jitter And Shimmer’, *IEEE International Conference on Computer, Communication, and Signal Processing*, Pages 1-6

Verde L, De Pietro G, Alrashoud M, Ghoneim A, Al-Mutib K.N, Sannino G, 2019, ‘Leveraging Artificial Intelligence to Improve Voice Disorder Identification Through the Use of a Reliable Mobile App’, *IEEE Access*, Pages 124048-124054

Critical Evaluation of object recognition models and Their Speed and Accuracy on Autonomous Vehicles

Luis Castrillo

Abstract

This paper provides an evaluation of different object detection methods focused in improve the accuracy and speed of previous models. This paper evaluates One-stage and Two-Stage object detection methods to compare the strengths and weaknesses of different models. Finally, the paper will conclude what will be the best model to enhance the accuracy and speed of object detection in autonomous car with RGB images.

1 Introduction

Into the computer vision technology object recognition models using machine learning (ML) is gaining considerable research interest. Object recognition is one of main objectives into different applications as autonomous driving, self-localization and mapping, real-time Environmental (Attila Börcs, et al., 2017).

Attila Börcs,et. al. (2017) explore the possibility to make the object detection with a Lidar. This research was made to create a laser-based technology capable of detect different type of objects in indoor environments using a feature sharing manner to use less amount of laser scanning data. They get the conclusion that this technology can be directly applied to 2-D object detection tasks.

Object detection with LiDAR sensor is more accurate than only use RGB images but LiDAR sensors are more expensive than RGB cameras. Zoltan Rozsa, Tamas Sziranyi (2019) proposed an object detection model using LiDAR that can detect and track objects obtaining more than 99% of accuracy in detection and tracking.

Juan M. Gandarias, et. al. (2019) compared three different methods to object detection using high-resolution tactile sensors. They compared the first method using only tactile images but this method has the limitation of the object complexity and the sensor quality. The second method used transfer learning that consist in a CNN trained with a large amount of RGB images and used to classify the pressure images. And the third method that consist in create a CNN trained with pressure image dataset. This third method achieved the most accuracy on object detection close to 100%.

This paper evaluates current research focused on enhanced the speed and accuracy of object recognition using RGB images and analysing these images with ML algorithms such as convolutional neural networks (CNN).

2 Current Object Detection Methods Using RGB Images

Majority object detection methods using ML have common network structure and use similar methodology. But based on current object detection literature methods that extract features from RGB images have different approaches.

This section is divided into two parts (2.1) present the features and methodology that One-stage methods have (2.2) present the features and methodology of Two-Stage methods.

2.1 One-Stage Object Detection methods

Detection of Regions of Interest (ROI) is the first stage in traditional object detection model. ROI is divided in two steps, extraction of features and achieve an accurate detection. But it requires a long detection time and cannot meet practical requirements.

This manner of object detection is called Two-Stage Object Detection. One stage object detection does not require the generation of ROI. The speed on detection increase but the accuracy is not high (Yijing Wang et. al., 2019).

According with the Yinjing Wang, et al (2019) research, they proposed a model that use One stage object detection network developing a BackBone network.

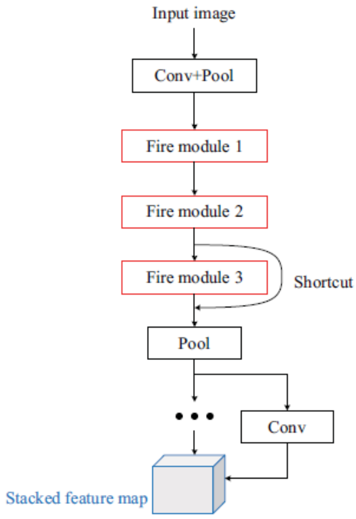


Fig. 1 BoneBack Network (Yijing Wang et. al., 2019).

They used SqueezeNet network with modules called Fire Module that squeeze and expand the image. They design a BackBone network shown in Fig. 1 that is improved replacing the traditional convolutional layer with the fire modules shown in Fig.2.

The first step is the generation of bounding boxes for each element into the image, they are generated by the Anchors Generation Network (AGnet). Next stage is the post-processing where the excess of anchors is eliminated based on the probability of the anchor contain a car and Intersection over Union (IoU) of the bounding box and the ground truth.

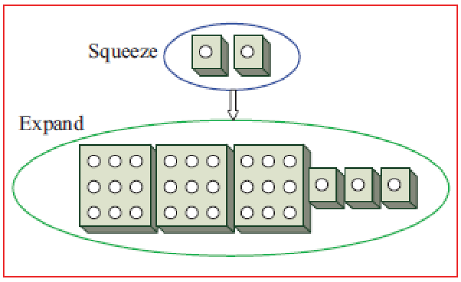


Fig. 2 Fire Module Structure (Yijing Wang et. al., 2019).

For the experiment authors used the KITTI dataset. They only considered cars for training and testing. In order to avoid the overfitting, they used different techniques as flipping, or cropping.

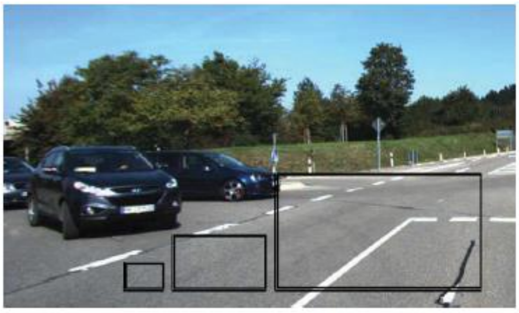
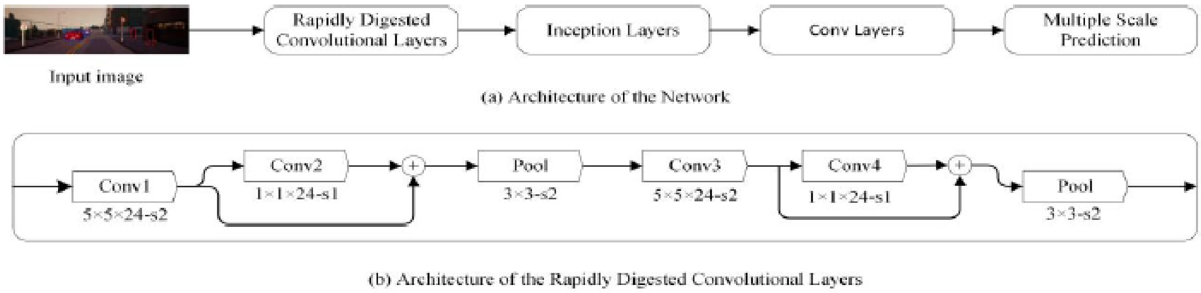


Fig. 3 Anchors of AGnet (Yijing Wang et. al., 2019).

Yijing Wang et. al. (2019) claims a one stage object detector specialized for car detection. And how Table 1 shown results of their proposed method improve previous researches and obtain in terms of speed 45.5 FPS which meets with real-time requirements. They claim to create a small model that means that is more suitable for deployment in autonomous vehicles (AV) systems.

The experiments conducted by Yijing Wang et. al. (2019) use KITTI dataset, it is a dataset used in the most object detection experiments that avoid bias and permit better comparation with other models, therefore claims are valid.

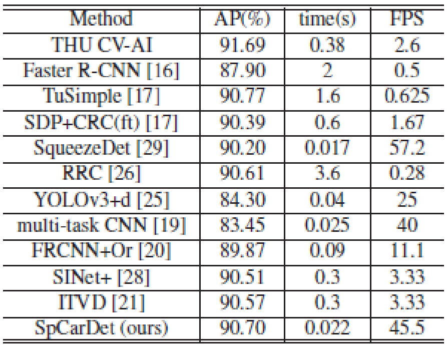


(a) Architecture of the Network

(b) Architecture of the Rapidly Digested Convolutional Layers

Fig. 4. Network Architecture (Zekun Luo, et al. 2019)

The results achieve the claim of the authors obtaining over 90% of object detection with low response time and with high speed that permit apply this method to the real autonomous cars. But the size of the proposed model is not clear because they said only that the model is small but not the size of it.



Method	AP(%)	time(s)	FPS
THU CV-AI	91.69	0.38	2.6
Faster R-CNN [16]	87.90	2	0.5
TuSimple [17]	90.77	1.6	0.625
SDP+CRC(ft) [17]	90.39	0.6	1.67
SqueezeDet [29]	90.20	0.017	57.2
RRC [26]	90.61	3.6	0.28
YOLOv3+d [25]	84.30	0.04	25
multi-task CNN [19]	83.45	0.025	40
FRCNN+Or [20]	89.87	0.09	11.1
SINet+ [28]	90.51	0.3	3.33
ITVD [21]	90.57	0.3	3.33
SpCarDet (ours)	90.70	0.022	45.5

Table 1: Test results (Yijing Wang et. al., 2019).

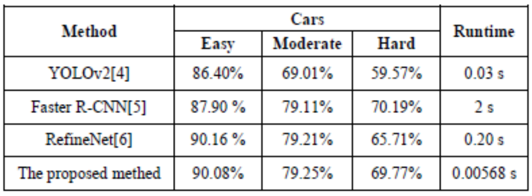
Zekun Luo, et al. (2019) proposed a one- stage object detection method applicable in real-time, detecting on-board object using a CNN on RGB-D images.

The first step is obtained RGB images with a binocular camera and calculate the objects' depth. Finally, the C.RELU is replaced with a 1-dimension convolutional layer to reduce the number of outputs as Fig.4(a) shown.

The proposed method by Zekun Luo, et al. (2019) achieved the results shown in Table 2, tested with the KITTI dataset. The most notorious improvement is in the runtime. Also, their model reaches a speed of 180fps. This allows to this method to have a better application in AV.

In this research paper Zekun Luo, et al. (2019) proposed a method that improve the effectiveness of object detection images in RGB-D images applicable in real-time, with high accuracy and with low-cost detection for AV systems.

The experiments conducted by Zekun Luo, et al. (2019) test their model with KITTI the dataset, used also to evaluate other method. They compare their results with the other systems results to check the improvements. Experiments are valid because they compare in the same environment their model with previous models .



Method	Cars			Runtime
	Easy	Moderate	Hard	
YOLOv2[4]	86.40%	69.01%	59.57%	0.03 s
Faster R-CNN[5]	87.90 %	79.11%	70.19%	2 s
RefineNet[6]	90.16 %	79.21%	65.71%	0.20 s
The proposed methed	90.08%	79.25%	69.77%	0.00568 s

Table 2: Test results (by Zekun Luo, et al., 2019).

They demonstrate that their model is faster than previous models such as YOLO or Faster R-CNN and with the speed obtained the model can be applied in real AV, therefore their claims are fully justified

Cheng-Yang Fu, et. al. (2017) proposed a single shot multibox detector (SSD) proposed by W. Liu, et. al. (2016) adding a de-convolutional layer to their proposed method in order to made a more effective model.

The SSD model use a CNN base network and add a series of smaller convolutional layers as Fig. 5 shown in blue color. These layers added are used for the prediction of scores and offsets for some predefine default bounding boxes.

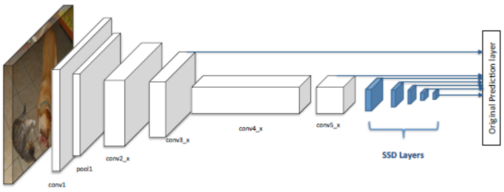


Fig. 5. SSD Network (Cheng-Yang Fu, et. al. 2017)

Cheng-Yang Fu, et. al. (2017) add de-convolutional layers as Fig. 6 shown in color red, in order to increase the resolution of the feature map. This model has to be pre-trained to take the advantage of transfer-learning for the decoder layers.

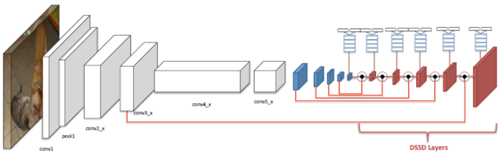
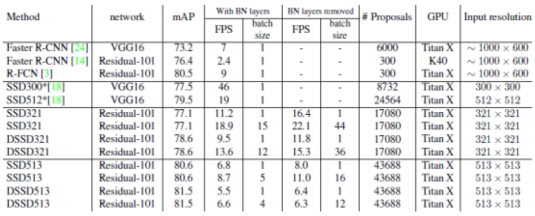


Fig. 6. DSSD Network (Cheng-Yang Fu, et. al. 2017)

To evaluate the speed and accuracy of their model they use COCO and PASCAL VOC2007 test that consist in a dataset with different pictures used to test the detection of multiple type of objects.

The experiments conclude that the new model is not faster because the addition of new layers made the process slower but they improve the accuracy of the previous model such as Faster R-CNN or SSD.



Method	network	mAP	With BN layers		BN layers removed		# Proposals	GPU	Input resolution
			FPS	batch size	FPS	batch size			
Faster R-CNN [24]	VGG16	73.2	7	1	-	-	6000	Titan X	∼ 1000 × 600
Faster R-CNN [14]	Residual-101	76.4	2.4	1	-	-	300	K40	∼ 1000 × 600
R-FCN [3]	Residual-101	80.5	9	1	-	-	300	Titan X	∼ 1000 × 600
SSD300*[18]	VGG16	77.5	46	1	-	-	8732	Titan X	300 × 300
SSD512*[18]	VGG16	79.5	19	1	-	-	24564	Titan X	512 × 512
SSD321	Residual-101	77.1	11.2	1	16.4	1	17080	Titan X	321 × 321
SSD321	Residual-101	77.1	18.9	15	22.1	44	17080	Titan X	321 × 321
DSSD321	Residual-101	78.6	9.5	1	11.8	1	17080	Titan X	321 × 321
DSSD321	Residual-101	78.6	13.6	12	15.3	36	17080	Titan X	321 × 321
SSD513	Residual-101	80.6	6.8	1	8.0	1	43688	Titan X	513 × 513
SSD513	Residual-101	80.6	8.7	5	11.0	16	43688	Titan X	513 × 513
DSSD513	Residual-101	81.5	5.5	1	6.4	1	43688	Titan X	513 × 513
DSSD513	Residual-101	81.5	6.6	4	6.3	12	43688	Titan X	513 × 513

Table 3: Test results (Cheng-Yang Fu, et. al. 2017).

Authors conclude that the new proposed model improve the detection of small objects against the SSD network and the relationship between these objects can be detected with more effectiveness.

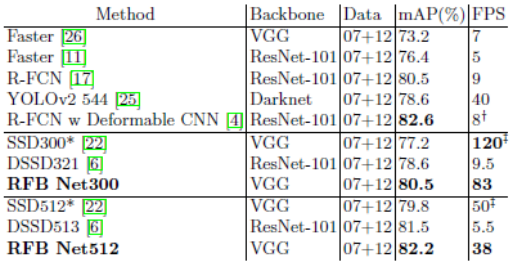
The experiments conducted used two datasets previously used to test the SSD model what means that the results can be compared in order to detect the improvements of the new model and in which aspects the new model is worse, therefore the experiments are valid.

They demonstrate that the new model is more robust and achieve better accuracy that the SSD method that means that the claim is valid.

Songtao Liu, et. al (2018) proposed a new object detection method named Receptive Field Block (RFB) based to made more strength the deep features of lightweight CNN what contribute to made faster and more accurate detectors.

Authors proposed a one-stage method that use a hand-crafted mechanism, RFB which imitates the structure of Receptive Field in human visual systems. This mechanism generates more robust and discriminative features.

For the experiments they use COCO and PASCAL VOC 2007 datasets. The use of these datasets enables the option of compare the method with the previous model in order to detect the improvements of the proposed model.



Method	Backbone	Data	mAP(%)	FPS
Faster [26]	VGG	07+12	73.2	7
Faster [11]	ResNet-101	07+12	76.4	5
R-FCN [17]	ResNet-101	07+12	80.5	9
YOLOv2 544 [25]	Darknet	07+12	78.6	40
R-FCN w Deformable CNN [4]	ResNet-101	07+12	**82.6**	$8^{\dagger}$
SSD300* [22]	VGG	07+12	77.2	**120**‡
DSSD321 [6]	ResNet-101	07+12	78.6	9.5
RFB Net300	VGG	07+12	**80.5**	**83**
SSD512* [22]	VGG	07+12	79.8	$50^{\ddagger}$
DSSD513 [6]	ResNet-101	07+12	81.5	5.5
RFB Net512	VGG	07+12	**82.2**	**38**

Table 4: Test results (Songtao Liu, et. al., 2018)

The results using PASCAL VOC 2007 dataset as Table 4 shown, achieve better results than previous improving the FPS and the accuracy. The proposed model gets more than 80 % of accuracy and enough speed to applicate the network in real world.

Authors of this paper use for the experiments COCO and PASCAL VOC 2007 dataset that avoid bias in the experiments and permit the fair comparation with the other models, therefore the experiments are valid.

Songtao Liu, et. al., (2018) proposed a method that improve the accuracy of the compared models and the speed of this method enable the possibility of use this mode in application in the real-world.

Firstly, to compare the accuracy of One-Stage object detection methods that use KITTI dataset to test the performance of the model, Yijing Wang et. al., (2019) achieve the best results in the object detection achieving a 90.7% of accuracy.

Comparing the speed of these models the new object detection prosed by Zekun Luo, et al. (2019) achieve the speediest method achieving a speed of 180 FPS. The speed achieved of this method and the accuracy demonstrate that both methods are suitable for use in the real AV systems.

Secondly for the comparation of One-Stage methods that use COCO and PASCAL VOC2007 datasets to test the performance of their methods, the model proposed by Songtao Liu, et. al., (2018) achieved the best accuracy with 82.2%.

2.2 Two-Stage Object Detection methods

The research proposed by Iljoo Baek, et. al. (2018) present an object detection model that detects in real-time objects using 4 fisheye cameras situated in a vehicle as Fig. 7 shown.

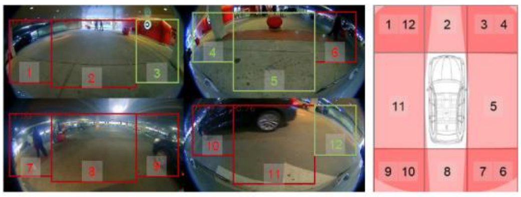


Fig. 7. "Mapping of the ROIs to regions around the vehicle" (Iljoo Baek, et. al. 2018)

The first element of this method is Detection of Moving objects as Fig.8 shown. This module is for detect if there is some object approaching from the left, right or center of each field of view.

The second element into the architecture is the tracking of stopped objects as Fig. 8 shown. This module solves the problem that if one moving object enters in a region of interest (ROI) and then stop the movement the previous module cannot detect it.

Finally, the object classification element as Fig. 8 shown is the module that classify ROI elements into different categories as pedestrians, vehicles, bicycles, etc.

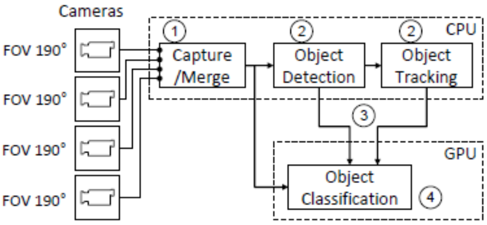


Fig.8: System Architecture (Iljoo Baek, et. al., 2018)

For the evaluation Iljoo Baek, et. al., (2018) used their model as the backbone of the R-CNN for fair comparison. They train the model using KITTI dataset and measured the efficiency using their own dataset. The system achieved 60.5% of accuracy recall and 81.1% of accuracy precision and 15.5 fps.

They claim a method that can detect, track and recognize moving object in real-time along previous regions of interest defined. The realized that the proposed model can be applied to any existing autonomous vehicle solution.

Authors train the model with the KITTI dataset but they measured the performance with their own dataset. That means that results cannot be compared with other systems that test their performance with the KITTI dataset, therefore the experiments of this method are not possible to compare.

Authors do not find a system that can be used in all the possible situations of autonomous cars. Their system gets poor results in the speed achieving only 15.5fps, therefore the conclusion of this research is not completely valid.

Fahimeh Farahnakian, et. al. (2018) proposed a Two- stage method capable of detect and identify object around AV in a maritime environment.

The first step of this method is the ROI extraction. To generate the initial proposal it generates the first initial proposal combining superpixels in an RGB image.

The bounding boxes generated are mapped on RGB image for filter the initial proposal.as Fig.9 shown the first ROI proposal have many Bounding boxes that do not contain any object, but after the filter the system just have the bounding boxes that contain an object.

The second step on the Fahimeh Farahnakian, et. al. (2018) method is the object classification. For this step, they developed a CNN as Fig.9 shows.

For the evaluation, Fahimeh Farahnakian, et. al. (2018) obtained the dataset collecting RGB images with two cameras and more data from an IR camera, LiDAR and radar in various weather conditions.

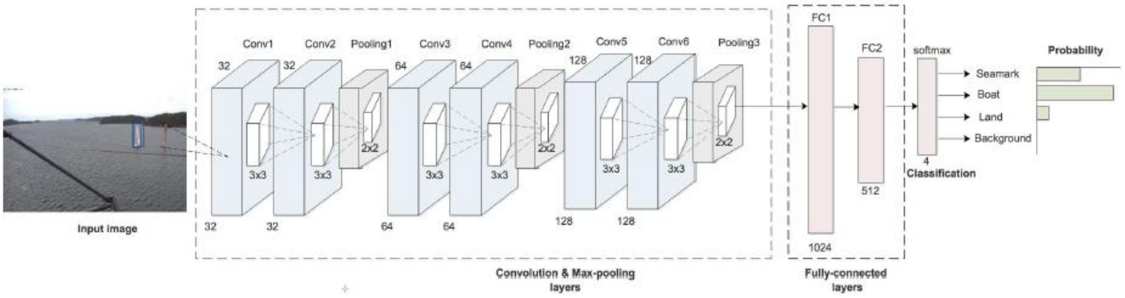
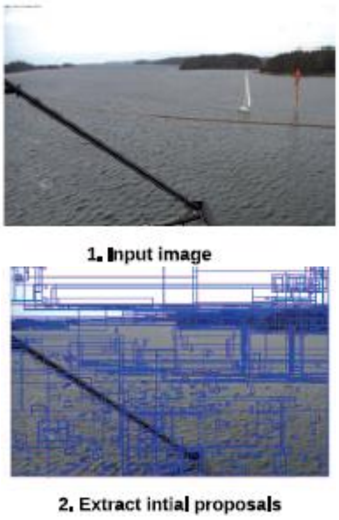


Fig.10 Proposed architecture. (Fahimeh Farahnakian, et. al. 2018)

The first stage is the CNN training, they used images from the RGB cameras. To detect object the extract local gradients groups that differed from the typical water surface. They train the CNN model with the dataset obtained previously which have 13,088 images.

The framework was evaluated with the real time dataset obtained before. They test the system using different CNN structures and they conclude that using the structure shown in Fig. 10 the obtained the best results.



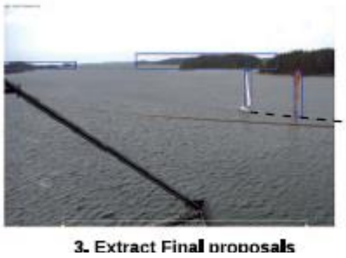
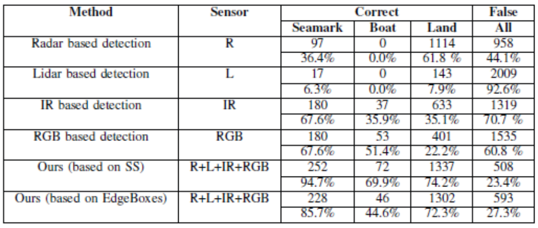


Fig. 9. (Fahimeh Farahnakian, et. al. 2018)

Experiment results as Table 5 shown, conclude that the proposed method can achieved better results detecting seamark, boats and land than the detection using the sensors by separated.



Method	Sensor	Correct			False
		Seamark	Boat	Land	All
Radar based detection	R	97	0	1114	958
		36.4%	0.0%	61.8 %	44.1%
Lidar based detection	L	17	0	143	2009
		6.3%	0.0%	7.9%	92.6%
IR based detection	IR	180	37	633	1319
		67.6%	35.9%	35.1%	70.7 %
RGB based detection	RGB	180	53	401	1535
		67.6%	51.4%	22.2%	60.8 %
Ours (based on SS)	R+L+IR+RGB	252	72	1337	508
		94.7%	69.9%	74.2%	23.4%
Ours (based on EdgeBoxes)	R+L+IR+RGB	228	46	1302	593
		85.7%	44.6%	72.3%	27.3%

Table 5. (Fahimeh Farahnakian, et. al. 2018)

Fahimeh Farahnakian, et. al. (2018) claim with this research an efficient object detection method that can be used by AV in maritime environment, reducing the computational cost and obtaining high accuracy in the localization and identification of objects using the region proposal method Selective search developed by . R. Uijlings, et. al.(2013);

Authors use a dataset created obtained specifically for this research and they demonstrate the improvement in the accuracy of their method compared with the object detection of each sensor separated, therefore experiments of this model are valid.

Authors demonstrate with their experiments how this method improve the accuracy of object detection and reduce significantly the objects no detected

Florian Chabot, et. al. (2017) proposed MANTA (Deep Many-Tasks), an object detection method capable of detect simultaneous vehicles, 3D dimensions and their parts in images even if these parts are not visible.

In this proposed method the divide the object analysis in two categories: 2D object detection and 3D object detection and pose estimation.

For the 2D object detection they used object proposal method, it proposed several boxes with high objectness confidence score. This region proposal method is combined with a CNN to refine the proposed regions of interest.

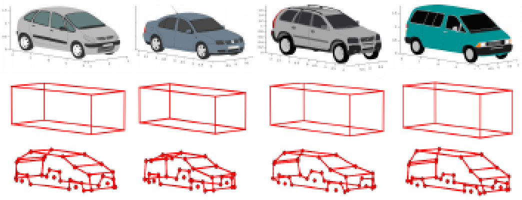
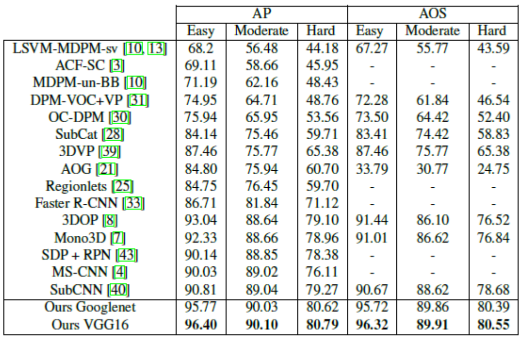


Fig.11 (Florian Chabot, et. al., 2017)

For the 3D pose estimation Florian Chabot, et. al. (2017) used 3D templates as Fig. 11 shown. Passing the 2D object detection stage output and the 3D templates a CNN predict the pose of the object, and the orientation. Finally, the pose estimation is made by 2D/3D vehicle part matching.

For the evaluation of their method Florian Chabot, et. al. (2017) use KITTI dataset. As Table 6 shown their proposed method improve the accuracy in 2D vehicle detection (AP) and orientation (AOS).

Florian Chabot, et. al., (2017) claim a new approach for 2D and 3D vehicle analysis from images. They generate vehicle bounding boxes using refinement steps and the 3D pose of the vehicle, even if these parts are hidden.



	AP			AOS		
	Easy	Moderate	Hard	Easy	Moderate	Hard
LSVM-MDPM-sv [10, 13]	68.2	56.48	44.18	67.27	55.77	43.59
ACF-SC [3]	69.11	58.66	45.95	-	-	-
MDPM-un-BB [10]	71.19	62.16	48.43	-	-	-
DPM-VOC+VP [31]	74.95	64.71	48.76	72.28	61.84	46.54
OC-DPM [30]	75.94	65.95	53.56	73.50	64.42	52.40
SubCat [28]	84.14	75.46	59.71	83.41	74.42	58.83
3DVP [39]	87.46	75.77	65.38	87.46	75.77	65.38
AOG [21]	84.80	75.94	60.70	33.79	30.77	24.75
Regionlets [25]	84.75	76.45	59.70	-	-	-
Faster R-CNN [33]	86.71	81.84	71.12	-	-	-
3DOP [8]	93.04	88.64	79.10	91.44	86.10	76.52
Mono3D [7]	92.33	88.66	78.96	91.01	86.62	76.84
SDP + RPN [43]	90.14	88.85	78.38	-	-	-
MS-CNN [4]	90.03	89.02	76.11	-	-	-
SubCNN [40]	90.81	89.04	79.27	90.67	88.62	78.68
Ours Googlenet	95.77	90.03	80.62	95.72	89.86	80.39
Ours VGG16	**96.40**	**90.10**	**80.79**	**96.32**	**89.91**	**80.55**

Table 6. Results for 2D vehicle detection(Florian Chabot, et. al., 2017)

For experiment and evaluate Florian Chabot, et. al., (2017) model they use KITTI dataset as the other systems to train and test the model and the comparation in the results achieve better results than previous proposed models, therefore experiments and evaluation are valid.

Authors created a model that can detect vehicles in 2D and 3D images for detection of vehicles and their pose, therefore they claim are justify.

The comparation of the result of models that use Two-Stage networks and KITTI dataset the system proposed by Florian Chabot, et. al.,(2017) achieve the best results obtaining a 10% more than the model proposed by Iljoo Baek, et. al., (2018).

The model proposed by Fahimeh Farahnakian, et. al. (2018) test their model with their own dataset, these experiments achieve results that demonstrate this model can be use in the real-world but cannot be compare with the other performance results.

3 Conclusions

Seeing all the literature reviewed in this paper, there are two main object recognition models, one-stage and two-stage. After all the analysis, the conclusion is that one-stage models fit better for AV application because all models get better results in real-time object detection, but they are less trustable in accuracy terms.

Two-stage methods make models more robust and precise and the precision for AV is very important because these vehicles will make their decisions based on the results that the model gives. Thus, object detection model for autonomous vehicle should have a perfect precision in their decisions but they are not suitable for AV application because the speed and real-time analysis is one crucial factor.

Taking all conclusions together, a good approach of a fast and precise object detection model can be developed a model merging the model proposed by Zekun Luo, et al. (2019) and Florian Chabot, et. al., (2017) where the model can achieve a very fast speed and strong accuracy making the best model. Therefore, more research is needed in this field to discover the way to develop the network that merge both models in a unique network.

References

Attila Börcs, Balázs Nagy, Csaba Benedek, 2017, ‘Instant Object Detection in Lidar Point Clouds’, *IEEE Geoscience and Remote Sensing Letters,* VOL. 14.

Cheng-Yang Fu, Wei Liu, Ananth Ranga, Ambrish Tyagi, Alexander C. Berg, 2017, ‘Dssd: Deconvolutional single shot detector’, In *arXiv*:1701.06659.

Fahimeh Farahnakian, Mohammad-Hashem Haghbayan, Jonne Poikonen, Markus Laurinen, Paavo Nevalainen, Jukka Heikkonen, 2018, ‘Object Detection based on Multi-sensor Proposal Fusion in Maritime Environment’, *17th IEEE International Conference on Machine Learning and Applications,* pp. 971 – 976.

Floria Chabot, Mohamed Chaouch, Jaonary Rabarisoa, Celine Teuliere, Thierry Chateau, 2017, ‘Deep MANTA: A coarse-to-fine many-task network for joint 2D and 3D vehicle analysis from monocular image’, *IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR),* pp. 1827–1836.

Iljoo Baek, Albert Davies, Geng Yan, Ragunathan Rajkumar, 2018, ‘Real-time Detection, Tracking, and Classification of Moving and Stationary Objects using Multiple Fisheye Images’, *IEEE Intelligent Vehicles Symposium (IV),* pp. 447 - 452.

Juan M. Gandarias, Alfonso J. García-Cerezo, Jesús M. Gómez-de-Gabriel, 2019, ‘CNN-Based Methods for Object Recognition with High-Resolution Tactile Sensors’, IEEE Sensors Journal, vol. 19.

R. Uijlings, K. E. Sande, T. Gevers, A. W. Smeulders, 2013, ‘Selective search for object recognition’, *Int. J. Comput. Vision,* 104(2):154–171.

S. Liu and D. Huang, 2018, ‘Receptive field block net for accurate and fast object detection", *European Conference on Computer Vision,* pp. 385-400.

W. Liu, D. Anguelov, D. Erhan, S. Christian, S. Reed, C.-Y. Fu, A. C. Berg, 2016, ‘SSD: single shot multibox detector’, *Lecture Notes in Computer Science journal,* pp. 21–37.

Yijing Wang, Wentao Sun, Zhiqiang Zuo, 2019, ‘Specialized Car Detector for Autonomous Driving’, *Proceedings of the 38th Chinese Control Conference,* pp. 6802 – 6807.

Zekun Luo, Xia Wu, Chunfu Luo, Ping Wang, 2019, ‘Real-time Object Detection Based on RGB-D Image’, *International Conference on Computational Electromagnetics (ICCEM),* Page 1-3.

Zoltan Rozsa, Tamas Sziranyi, 2019, ‘Object Detection From a Few LIDAR Scanning Planes’, *IEEE Transactions on Intelligent Vehicles,* Vol.4., pp. 548 – 560.

A Critical Evaluation of Artificial Intelligence Research Aimed At Improving Robotic System To Help Visually Impaired People

Oussama El Jerche

Abstract

As a result of its huge expansion, technology aids visually impaired people interact and communicate with their surroundings and several techniques are now available for providing guidance in so many ways. This paper compared and evaluated some methods used to improve robotic systems, these include: Fuzzy Logic Control, Visual Odometry and Place Recognition to help visually impaired people. An evaluation of some of these methods has been discussed and comparisons of techniques are carried out through the paper to figure out the suitable method, and recommendations are made for some future works.

1 Introduction

Visually impaired people are often unaware of dangers in front of them. The challenge with current systems is that the problem area is becoming more and more complex. They are lacking to solve navigation problems, and this incorporates discovering objects, accurately putting things, and distinguishing various colors, content, or other visual patterns.

Patel and Chinchole (2018) provided that some electronic aids can provide a lot of information to the user about his environment to improve his daily life. Toha S.F. et al, (2015) argued that visually impaired people need path guidance robot and they cannot rely on their sense of hearing because it’s going to put them in danger, so that they used an ipath system which includes a PID controller to give a quick response to the visually impaired person to take his route.

In a study to find out the shortest path possible for visually impaired people, Krishnan (2009) proposed a robot path planning method using the Ant algorithm and showed that the algorithm uses its own secretions to find the shortest route.

The research conducted by Kaiser and Lawo (2012) in mapping and the situation of walkers during the investigation of the obscure conditions suggest choosing and find the best routes possible effectively to guide visually impaired persons.

This research paper gives a critical evaluation of the current techniques used to guide visually impaired people which are Fuzzy Logic Control, Visual Odometry and Place Recognition, and it is organized as follows: outlining and evaluating the current methods used to guide visually impaired people. Providing the comparisons of the discussed methods and the conclusions on the research paper are reported in the end.

2 Guiding robot’s methods

This section presents three techniques used to guide visually impaired people. These included: Fuzzy Logic Control, Visual Odometry and Place Recognition. It will examine the methods and the validity of the experiments as well as the ramifications of the outcomes.

2.1 Fuzzy Logic

Wei Y. et al, (2013) proposed a method known as the Fuzzy Logic Control to recognize little automatic power control and the normal navigational advancement of the user, in order to guide visually impaired people in indoor environments. The study was directed on some visually impaired people to communicate with a smart rope system which contains one lobby sensor joystick so it can lessen the control blunders and ameliorate the preparing speed. Individuals gave reactions reliant on the power sense.

Wei Y. et al, (2013) chose fuzzy logic as an intelligent classifier to determine movement control power and startling movement. They planned two fuzzy controllers and powers were isolated by recurrence and they compared the method with their previous one to decide if there is an improvement of one on the other.

The parameters for handling are course and the movement changing recurrence. Figure 1 and Figure 2 represent the fuzzification for both parameters, also the result conduct of the sign being gotten by the fuzzy controller and these ensuing standards can be found in Figure 3.

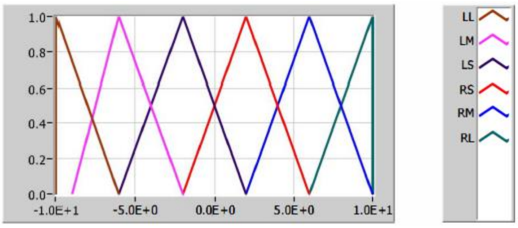


Figure 1 Right motor direction membership functions. Wei Y. et al, (2013)

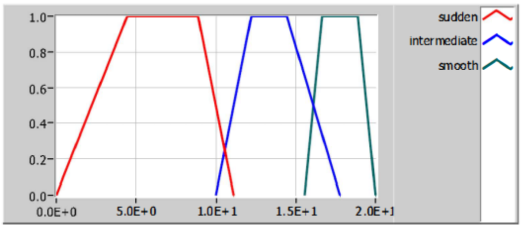


Figure 2 Membership functions of right motor's motion changing frequency. Wei Y. et al, (2013)

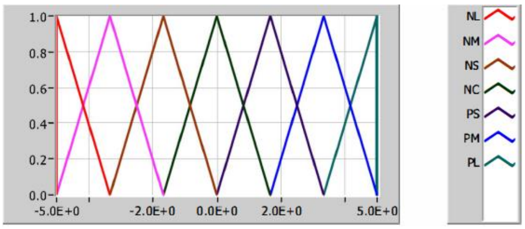
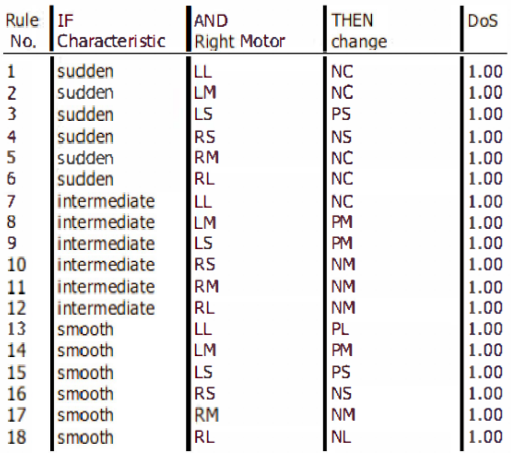


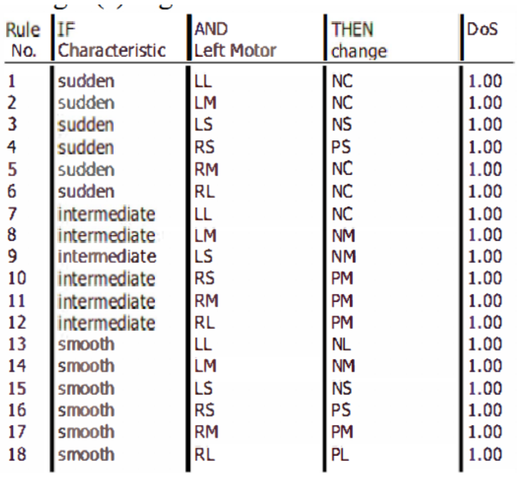
Figure 3 Output consequent fuzzy rules for fuzzy controller. Wei Y. et al, (2013)

The outcomes appeared in Figure 4 (an) and (b) use a method to align the Fuzzy principles in order to ensure the exact interaction. A defuzzification method is given to provide an express worth result of the robot's wheels and also to terminate the technical control.



Rule	IF	AND	THEN	DoS
No.	Characteristic	Right Motor	change	
1	sudden	LL	NC	1.00
2	sudden	LM	NC	1.00
3	sudden	LS	PS	1.00
4	sudden	RS	NS	1.00
5	sudden	RM	NC	1.00
6	sudden	RL	NC	1.00
7	intermediate	LL	NC	1.00
8	intermediate	LM	PM	1.00
9	intermediate	LS	PM	1.00
10	intermediate	RS	NM	1.00
11	intermediate	RM	NM	1.00
12	intermediate	RL	NM	1.00
13	smooth	LL	PL	1.00
14	smooth	LM	PM	1.00
15	smooth	LS	PS	1.00
16	smooth	RS	NS	1.00
17	smooth	RM	NM	1.00
18	smooth	RL	NL	1.00

Figure 4 (a) Right wheel controller rule base. Wei Y. et al, (2013)



Rule	IF	AND	THEN	DoS
No.	Characteristic	Left Motor	change	
1	sudden	LL	NC	1.00
2	sudden	LM	NC	1.00
3	sudden	LS	NS	1.00
4	sudden	RS	PS	1.00
5	sudden	RM	NC	1.00
6	sudden	RL	NC	1.00
7	intermediate	LL	NC	1.00
8	intermediate	LM	NM	1.00
9	intermediate	LS	NM	1.00
10	intermediate	RS	PM	1.00
11	intermediate	RM	PM	1.00
12	intermediate	RL	PM	1.00
13	smooth	LL	NL	1.00
14	smooth	LM	NM	1.00
15	smooth	LS	NS	1.00
16	smooth	RS	PS	1.00
17	smooth	RM	PM	1.00
18	smooth	RL	PL	1.00

Figure 4 (b) Left wheel controller rule base. Wei Y. et al, (2013)

The conclusions drawn in this paper tend to be reasoned that the scientists demonstrated that this technique is powerful during driving, hindrance shirking and following for the framework's intelligent plan and it has the appropriateness to the dubious system with the obscure model.

Looking at this method, it tends to be seen that the two Artificial Intelligence procedures were utilized to control outwardly weakened individuals in the real world. All methodologies have clear destinations and were fittingly reported with the goal that different researchers might have the option to go over the analysis subsequently thus analyze and improve the findings. In order to guarantee exact communication, the entire fuzzy control plot was structured as a classifier. The research led by Wei Y. et al, (2013) was a well-thought approach.

There have been many other studies, such as Stephen (2000) and Alhmiedat et al, (2013) which give evidence of a strong correlation with collaborator gadget improvement for visually impaired people dependent on minimal effort devices.

2.2 Visual Odometry

Research carried by Nguyen Q.H. et al, (2014) used a visual odometry method dependent on the following of ground plane highlights to construct a course of the movement. They proposed to utilize Kalman Filter, which helps in combining coordinated results of the present perception and the estimation of robot states, to overcome the issue of a present perception that could be coordinated with an extremely far backward picture that makes inaccurate localization of the robot.

Nguyen Q.H. et al, (2014) evaluated the proposed technique in a passage situation of a structure, they uncovered execution on the adequacy of utilizing Kalman Filter in directing the robot to affirm their structure is plausible to stand up to useful issues in the developments of the robot. They proposed to address the drift with the utilization of image matching based restriction in combination with Kalman filter where robot moves seek after a straight street. Figure 5 portrays route information without and by utilizing the Kalman filter.

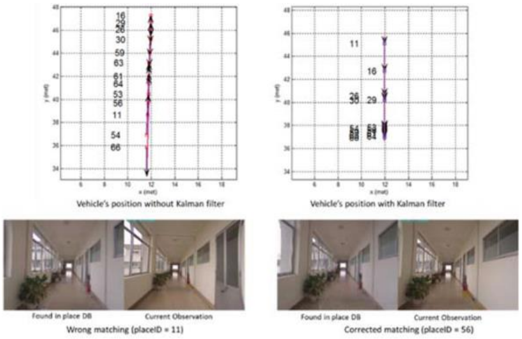
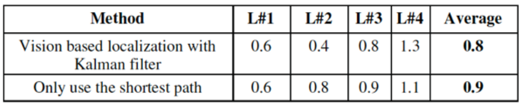


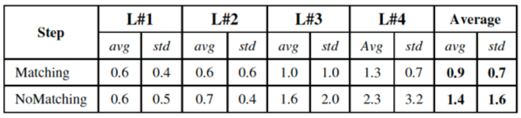
Figure 5 Vehicle moving without/with Kalman Filter. Nguyen QH. et al (2014)

Nguyen Q.H. et al, (2014) estimated middle and normal blunder of evaluated position to survey the restriction precision if there should be an occurrence of use/nonuse of Kalman Filter since utilizing just the spot acknowledgment results, the bearings supporting route administrations are clearly unrestrained. Tables 1 and 2 show the evaluation results.



Method	L#1	L#2	L#3	L#4	Average
Vision based localization with Kalman filter	0.6	0.4	0.8	1.3	**0.8**
Only use the shortest path	0.6	0.8	0.9	1.1	**0.9**

Table I. Median Error (in Meters). Nguyen Q.H. et al, (2014)



Step	L#1		L#2		L#3		L#4		Average	
	avg	*std*	*avg*	*std*	*avg*	*std*	*Avg*	*std*	*avg*	*std*
Matching	0.6	0.4	0.6	0.6	1.0	1.0	1.3	0.7	**0.9**	**0.7**
NoMatching	0.6	0.5	0.7	0.4	1.6	2.0	2.3	3.2	**1.4**	**1.6**

Table II. Average Error (in Meters). Nguyen Q.H. et al, (2014)

The scientists concluded that the method is feasibly navigating visually impaired people and provide deploying navigating services in indoor environments, although future research could look into similar techniques.

Looking at the method results presentation, it was fair, giving it the advantage of being scientifically sound. The majority of claims and conclusions were justified as the authors' supports with results that proved their claims. The experiments were not biased as there was a correlation of the new technique with various algorithms to quantify execution of its precision and sufficiency, along these lines making the method conceivable and its experiment can be repeatable. The research carried out by Nguyen Q.H. et al, (2014) was a thorough and well-thought approach.

However, the research led by Pradeep V. et al, (2010) gives a strong correlation between Visual Odometry and highlight based metric-topological SLAM (Simultaneous localization and mapping) for distinguishing hindrances in the way and alert subjects about the nearness of visually impaired people.

2.3 Place Recognition

In their research, Cummins and Newman (2008) proposed a Place Recognition method which utilizes a probabilistic approach called FAB-MAP that allows them to unequivocally represent perceptual aliasing in the environment and guide visually impaired people. They propelled the best in class by building up a principled probabilistic framework for appearance-based place recognition. To adjust the issue that the visual introductions should be simple execution and productive distinctive scenes, Cummins and Newman (2008) involved the FAB-MAP algorithms which made a good achievement in coordinating spots in directions over a long period of time.

In their framework, they provided an algorithm so as to deduce the likelihood of circulation and the connections between the visual words. The FAB-MAP incorporates co-happening visual expressions of a similar subject. Cummins and Newman (2008) tested the portrayed algorithm utilizing symbolism from a versatile robot. The latter gathers every picture and pass it into a preparing pipe that creates a pack of words portrayal.

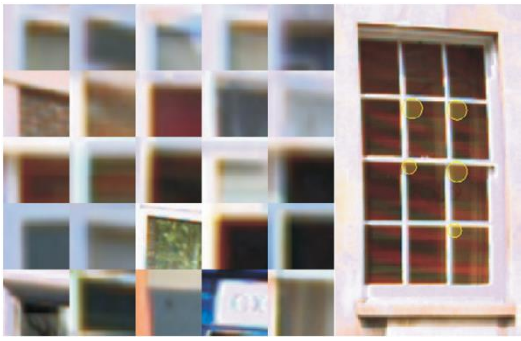


Figure 6. Interest points quantized to this word typically correspond to the top-left corner of windows. Cummins and Newman (2008)

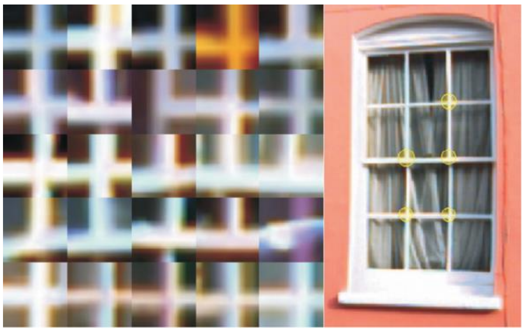


Figure 7. Interest points quantized to this word typically correspond to the cross-piece of windows. Cummins and Newman (2008)

The robot accumulates pictures to one side and right of its direction around each 1.5 m. Each assembled picture is set up by their algorithm and is utilized to instate another spot, if loop closure is identified, or to refresh a current spot model. Figure 8 depicts the evaluation results.

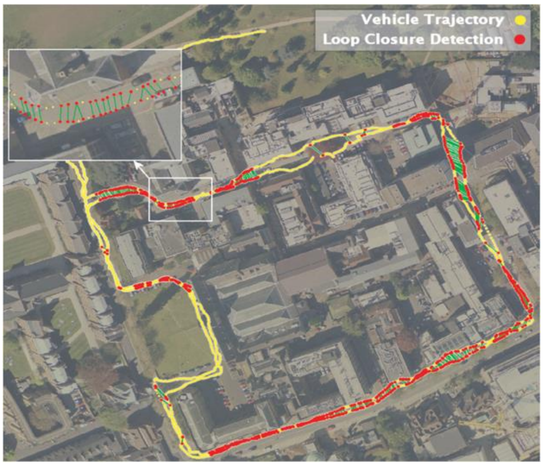


Figure 8 Appearance-based matching results for the City Centre data set overlaid on an aerial photograph. Cummins and Newman (2008)

It has been concluded that the researchers indicated that the Fast Appearance-Based Mapping is the best way to deal with place recognition and was effective in distinguishing huge parts of circle terminations in testing outside conditions, without bogus positives.

This research used the SURF (speeded up robust features) detector to remove districts of enthusiasm from pictures and compute 128D non-pivot invariant descriptors for these districts and it was achieved by clustering every one of the descriptors from a lot of preparing pictures utilizing a straightforward steady grouping technique. All images were collected in urban streets which helps to confirm that the method will be applicable in the real world. The research conducted by Cummins and Newman (2008) was a well thought and in-depth approach.

However, the research carried out by Chumkamon et al, (2008) provides a strong relation between place recognition and RFID to help the visually impaired individuals reach a path to a specified destination and locate places.

3 Comparisons of the methods

The research evaluated throughout this paper all had an aim to improve robotic systems via some useful techniques. The methods all had advantages in certain cases and limitations stated in their conclusions and results showing comparisons to be made for a recommendation for the most appropriate method.

The method proposed by Nguyen Q.H. et al, (2014) had advantages of localization accuracy as the results show that they had a lower error rate using Kalman Filter. This is a significant advantage as low localization accuracy will provide fewer problems for the visually impaired people when they would go travel in their environments.

The research conducted by Cummins and Newman (2008) used a probabilistic appearance-based approach to make their method function. FAB-MAP does also have a lower localization rate, although not as low as the Kalman Filter method. It does, however, have a review rate that is adequate to distinguish practically all circle terminations which will give success in recognizing enormous segments of circle terminations for

the visually impaired individuals in challenging outdoor environments. Thus, research into methods extending from Visual Odometry can improve the real-life application of this method.

On the other hand, the research by Wei Y. et al, (2013) showed that, whilst Fuzzy Control Logic had advantages in leading and obstacle avoidance, there were disadvantages in traffic detection which not ensure the visual impairer's safety. However, this simultaneous method is inferior to the other methods analysed as it shows no significant advantages in guiding visually impaired people.

The results underlined that all tested methods had limitations in one or multiples areas which included real-world application. Thus, requiring further research to be conducted to achieve fully intuitive results.

4. Conclusions

This paper evaluated different researches carried out by the different researchers to improve Robotic systems. These researches included Fuzzy Logic Control, Visual Odometry and Place Recognition. Upon critically evaluating the above-mentioned technologies, it clearly shows that the Visual Odometry method provided by Nguyen Q.H. et al, (2014) is the most suitable methodology to guide visually impaired individuals so that it fundamentally centers around neighborhood consistency and expects to gradually gauge the way of the camera/robot pose after pose and conceivably performing nearby improvement.

Cummins and Newman (2008) also provided a significant experiment into algorithms that can be successful and helpful in detecting big portions of circle terminations in outdoor environments, providing a detailed method, results and generative models. However, their experiment was flawed with a bit lack of testing, only using one type of algorithms.

The Fuzzy Logic Control method provided by Wei Y. et. al, (2013) indicated that further research into Artificial Intelligence for robotic systems could yield more accurate and thus, safer.

Generally, looking at all the methods mentioned above, they all brought good results. The fight towards eliminating the problem with guiding visually impaired people, in any case, the techniques above have demonstrated that it can be done, with the right work put in.

References

Alhmiedat T., Taleb A.A, Samaraz G., 2013, 'A prototype navigation system for guiding blind people indoors using NXT Mindstorms.', *International Journal of Online Engineering,* Vol. 9, Issue 5, Pages 52-58

Chumkamon S., Tuvaphanthaphiphat P., Keeratiwintakorn P., 2008, ‘A blind navigation system using RFID for indoor environments.’, *Proceedings of the 5th International Conference on Electrical Engineering/Electronics,* Pages 765–768.

Cummins M. and Newman P., 2008, ‘FAB-MAP: Probabilistic localization and mapping in the space of appearance.’, *International Journal of Robotics Research,* Vol. 26, Issue 6, Pages 647-665.

Kaiser E.B., Lawo M., 2012, ‘Wearable navigation system for the visually impaired and blind people.’, *Proceedings - IEEE/ACIS 11th International Conference on Computer and Information Science*, Issue 1, Pages 230-233.

Krishnan S.B., 2009, ‘Robotic Wheelchair to guide elderly and visually impaired people.’, *IEEE International Conference on Technologies for Practical Robot Applications,* Issue 2, Pages 23-28.

Nguyen Q.H., Vu H., Tran T.H., Nguyen Q.H., 2014, ‘A vision-based system supports mapping services for visually impaired people in indoor environments.’, *13th International Conference on Control Automation Robotics and Vision,* Pages 1518-1523.

Patel S., Chinchole S., 2018, ‘Artificial intelligence and sensors based assistive system for the visually impaired people.’, *Proceedings of the International Conference on Intelligent Sustainable Systems*, Pages 16-19.

Pradeep V., Medioni G., Weiland J., 2010, ‘Robot vision for the visually impaired.’, *proceeding of the 5th International Conference on Communications and Electronics,* Pages 15-22.

Stephen S.E., 2000, ‘Zebra-crossing detection for the partially sighted.’, *IEEE Conference on Computer Vision and Pattern Recognition,* Vol 2, Pages 211-217.

Toha S.F., Yusof H.M., Razali M.F., Halim A.H., 2015, ‘Intelligent path guidance robot for blind person assistance.’, *4th International Conference on Informatics, Electronics and Vision,* Pages1-5.

Wei Y., Kou X., Lee M., 2013, ‘Development of a guide-dog robot system for the visually impaired by using fuzzy logic-based human-robot interaction approach.’, *International Conference on Control, Automation and Systems,* Pages 136-141.

Critical Evaluation of Current Methods Aimed at Improving Ransomware Detection

Thomas Herb

Abstract

With the increasing rampage of the ransomware infections and the related economic damage, ransomware has drawn the attention of cyber security researchers all over the world. This paper evaluated current research in the area of static and dynamic ransomware detection which are the most commonly used approaches. The analysis and comparison of static and dynamic methods shows how accurate and sustainable methods currently perform. Thereby, the paper presents conclusions showing gaps in current knowledge about ransomware detection and recommends a solution being the most effective for computer systems to stay secure.

1 Introduction

Ransomware detection methods can be divided in two categories: static and dynamic methods whereby neither of them shows an accuracy of 100% under all circumstances (Bander Ali Saleh Al-rimy et. al., 2018).

Research is carried out on numerous areas like the detection of ransomware specifically for Android platforms like the proposal of Iram Bibil et. al. (2019).

Khaled Alrawashdeh and Carla Purdy (2018) proposed a hardware-based solution to detect ransomware which is suited for devices with limited hardware resources like IoT devices.

Aviad Cohen and Nir Nissem (2018) invented a technique aimed to detect ransomware specifically in cloud systems.

This paper presents an evaluation of software-based ransomware detection methods which can be applied by common computer systems without remarkable limitations in hardware resources. Due to the fact that previous research in this area is limited to isolated proposals, this paper serves the purpose to gain knowledge on the validity of static and dynamic proposals including their advantages and disadvantages in comparison to each other. Based on that, a recommendation will be made which solution is the best for computer systems to stay secure from ransomware.

2 Evaluation of Ransomware Detection Methods

This section presents an evaluation of current ransomware detection methods.

2.1 Evaluation of Static Based Methods for Ransomware Detection

The following section presents an evaluation of static based methods which detect ransomware by recognizing static characteristics typical for ransomware.

Krzysztof Cabaj et. al. (2017) proposed a method using the size of HTTP packages to detect ransomware. The method observes the network of a computer system to identify HTTP packages being sent to a Command & Control (C&C) server to download the encryption key. The typical procedure of a ransomware infection is presented in Figure 1 and it shows the size of the HTTP packages which is essential for the ransomware detection of this method.

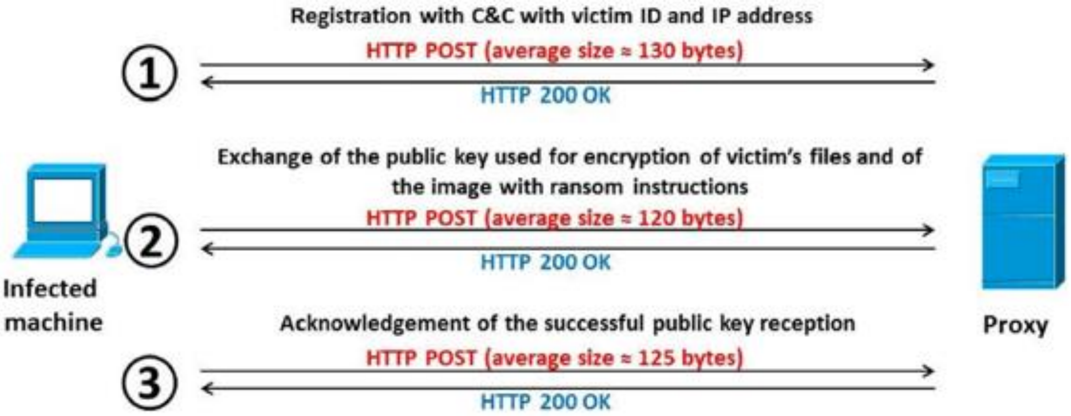

Registration with C&C with victim ID and IP address
HTTP POST (average size ≈ 130 bytes)
HTTP 200 OK
①
Exchange of the public key used for encryption of victim's files and of the image with ransom instructions
HTTP POST (average size ≈ 120 bytes)
HTTP 200 OK
②
Infected machine
Proxy
Acknowledgement of the successful public key reception
HTTP POST (average size ≈ 125 bytes)
HTTP 200 OK
③

Figure 1 CryptoWall Communication (Krzysztof Cabaj et. al., 2017)

The method was tested in a controlled experiment in an isolated environment with 787 samples of two ransomware families.

The method shows an overall accuracy of circa 97% with 2% false negatives.

According to the authors, that's an efficient method since HTTP characteristics are sufficient to detect ransomware reliably. Additionally, the analysis of the HTTP packages can help to identify and blacklist attackers' C&C servers.

The method was tested with samples of two ransomware families which makes the validity of the results questionable. The number of used samples was adequate but a larger variety of families is more essential since samples of the same family share the most characteristics. Therefore, it's probable that a static method shows a high accuracy within a collection of one ransomware family but it could perform worse on other families.

Nevertheless, the method could show similar results on other ransomwares what needs to be tested. A disadvantage is the fact that the size of HTTP packages could be altered by attackers making the method unable to detect them and the communication with C&C servers could be processed by HTTPS which would prevent the identification of C&C servers. Therefore, the method is no sustainable solution for the ransomware problem.

Another method using network traffic to detect ransomware was proposed by Daniel Morato et. al. (2018). Thereby, anomalies in SMB packages are recognized to detect ransomware attempts to encrypt data stored on NAS devices.

Experiments were carried out with 19 ransomware families and the method showed an accuracy of 100% before at most 10 files are encrypted and almost none false positives.

That means that the method is currently an almost optimal solution for the prevention of the encryption process on computer systems storing user data only on shared volumes.

The both methods of Krzysztof Cabaj et. al. (2017) and Daniel Morato et. al. (2018) could be useful for computer systems where computing power of client devices should not be used to detect ransomware. Thereby, the solution of Daniel Morato et. al. (2018) would only be the better solution for computer systems using NAS storages since it is unable to prevent ransomware from encrypting data on client devices.

The 100% accuracy of the method of Daniel Morato et. al. (2018) relies on the fact that the authors apparently found a characteristic which all current ransomware families share. If attackers change the SMB packages sent by ransomware it could stay undetected by the method. This is a realistic future issue the method could face and therefore, it is possible that it is no sustainable solution for the ransomware problem for computer systems using shared volumes.

May Medhat et. al. (2018) developed a framework which detects ransomware based on numerous features which were collected from samples belonging to 45 ransomware families. These features include cryptographic signatures, API functions and file keywords.

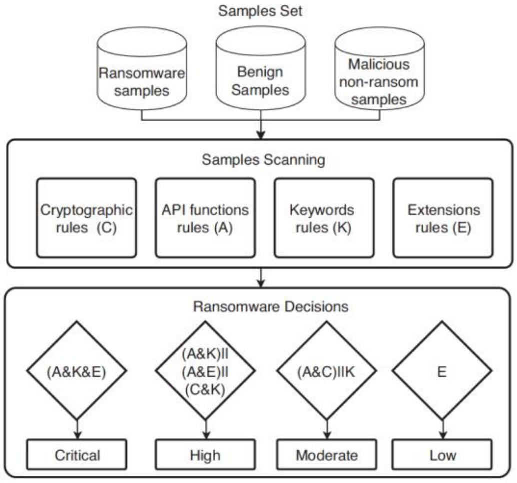


Figure 2 Proposed Framework (May Medhat et. al.,2018)

Based on the for ransomware typical features, decision rules are implemented on which incoming files are tested. Depending on the probability that the file could be ransomware an alarm occurs warning the user.

This research is based on 3 sets of samples: 793 ransomware files (45 families), 878 benign files, and 898 non-ransom malicious files.

The validity of the results was exemplary proven by the in-depth analysis of one ransomware sample and how it matches to the for ransomware typical features.

The results presented in Figure 4 show several accuracy values including an overall accuracy of 94,14%. The false positives rate is 8,4%. Moreover, 17 new ransomware families which were not included in the training set were detected.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

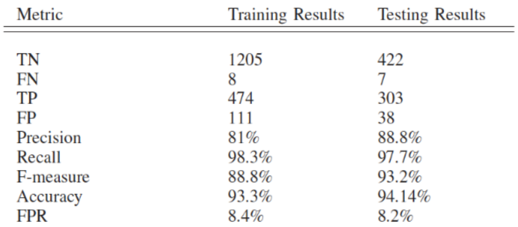
$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F-\ measure = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (3)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$FPR = \frac{FP}{FP + TN} \quad (5)$$

Figure 3 Framework Performance Metrics (May Medhat et. al.,2018)



Metric	Training Results	Testing Results
TN	1205	422
FN	8	7
TP	474	303
FP	111	38
Precision	81%	88.8%
Recall	98.3%	97.7%
F-measure	88.8%	93.2%
Accuracy	93.3%	94.14%
FPR	8.4%	8.2%

Figure 4 Training and Testing Results (May Medhat et. al.,2018)

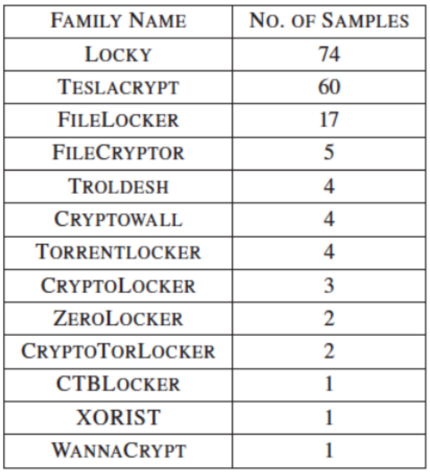
The authors claim that their approach to use decision rules based on multiple features of ransomware can help to improve the detection of ransomware since the results are very accurate. That is underlined with the comparison of their accuracy to commercial antivirus solutions.

The variety and number of used samples is vast which is a sign for the strong fundament of the results. However, it is not thoroughly mentioned how the experiments were carried out since it only presents a section with the analysis of one sample which is wrongly called a case study. That makes the results questionable and potentially biased since the experiments are not repeatable. The comparison to the accuracy to commercial antivirus solutions is questionable as it is not clear whether similar datasets were used.

Poudyal et. al. (2018) proposed a static method using machine learning algorithms to detect ransomware. Suspicious files are reverse engineered to assembly instructions and .dll files which are analyzed to find features being typical for ransomware.

The method was tested with 8 machine learning classifiers on 178 ransomware samples belonging to 13 ransomware families in an isolated environment.

As Figure 5 shows, the approach to use .dll files and Assembly instructions combined is on average the most accurate method with 95,5%.



FAMILY NAME	NO. OF SAMPLES
LOCKY	74
TESLACRYPT	60
FILELOCKER	17
FILECRYPTOR	5
TROLDESH	4
CRYPTOWALL	4
TORRENTLOCKER	4
CRYPTOLOCKER	3
ZEROLOCKER	2
CRYPTOTORLOCKER	2
CTBLOCKER	1
XORIST	1
WANNACRYPT	1

Figure 5 Ransomware Families (Poudyal et. al.,2018)

The point making this approach unique is according to the authors the fact that multiple aspects of ransomware are taken into consideration in order to detect it more reliably. Therefore, the accuracy can be improved while other solutions try to detect ransomware by only a single characteristic. This multi-level analysis of .dll files and assembly instructions should help to understand the behavior of ransomware better.

As Figure 5 shows, the vast majority of the samples used belong to only two ransomware families. Based on the results, the approach of the authors seems to be promising but tests need to be carried out with more samples and ransomware families.

The conclusions drawn by the authors are based on the good accuracy of their method. If the accuracy is confirmed by being tested with more samples this method will be a valuable approach to detect ransomware.

2.2 Evaluation of Dynamic Methods

This section shows an evaluation of dynamic approaches which detect ransomware by analysing it's behaviour like read- and write operations what is mostly conducted in isolated sandbox environments.

Tianliang Lu et. al. (2017) proposed a dynamic method detecting ransomware with the V-detector negative selection algorithm which is a learning algorithm based on an artificial human immune system. API calls, network operations and memory patterns are used to determine a behaviour which is specific for ransomware.

The method was tested in an isolated virtual machine containing dummy data being typically encrypted by ransomware. The ransomware sample was executed in a sandbox environment to test whether it would infect the host.

The experiment was carried out with 1000 samples belonging to at least 5 ransomware families whereby the exact number of families was not mentioned. One half of the samples was used to train the algorithm whereas the other half was used to test the detection accuracy.

The results vary due to the fact that the algorithm can be used differently depending on a variable called self-radius. The lower the value of the self radius the higher is the detection rate but also the rate of false alarms. This issue is presented in the Figures 6 + 7.

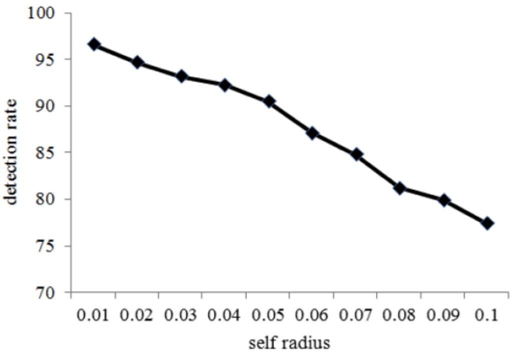


Figure 6 Detection Rate of Different self-radius (Tianliang Lu et. al.,2017)

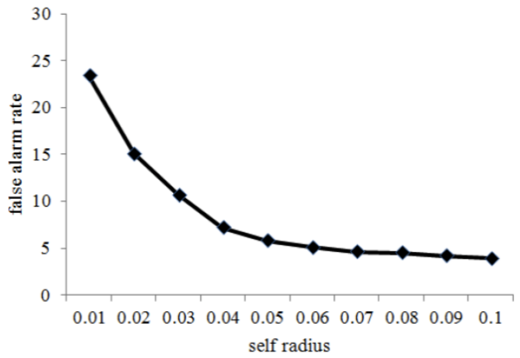


Figure 7 False Alarm Rate of Different self-radius (Tianliang Lu et. al.,2017)

According to the authors, a value between 0.04 and 0.05 is a good choice since the detection rate is over 90% and the false alarms rate below 10%.

The authors claim that the proposed method is a solution to especially detect novel variants of ransomware accurately.

The number of ransomware samples used is adequate to support the claim but without mentioning the exact number of used ransomware families the results remain questionable. Additionally, it is not clear whether the method detects novel ransomware samples since the separation of learn and test samples is not mentioned neither. Therefore, it is not stateable whether only the ransomware samples used to train the algorithm were detected.

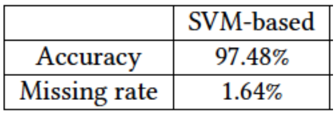
The false detection rate of almost 10% is high since it could prevent users from receiving many valid data disturbing the usage of the IT infrastructure in an enormous manner. It is not clearly stated whether this problem could be solved. Training the algorithm with additional samples from more ransomware families could improve the accuracy of the method what needs to be tested.

Yuki Takeuchi et. al. (2018) proposed a method detecting ransomware with support vector machines being a supervised machine learning algorithm. Thereby, API calls like read, write and create are analysed in order to determine behaviour of ransomware.

The experiments were conducted in a set up virtualized environment using a VM as a sandbox environment in which the samples are executed and analysed. The victim host is a windows VM including dummy data which is usually targeted by ransomware.

The method was tested with 272 samples of at least 4 different ransomware families whereby the exact number of families was not mentioned. 321 samples of benign software were used to test the number of false positives.

The results show an accuracy of 97.48% and a false negative rate of 1.64%.



	SVM-based
Accuracy	97.48%
Missing rate	1.64%

Figure 8 Accuracy and False Negative Rate (Yuki Takeuchi et. al.,2018)

Claimed by the authors, this method is unique because it inspects ransomware API calls more deeply helping to detect especially novel variants more accurately.

The claims of the authors are partly questionable because of the fact that it is not clear how many ransomware families were exactly used to test the method. Another point staying unclear is whether the method actually detected novel variants of ransomware because it is unknown which particular samples were detected. Leaving out the result of false positives makes it difficult to assess the method appropriately. A thorough description of the result as presented in Figure 3 would be needed.

However, the accuracy and false negative rate are excellent meaning that the approach of the authors could be a promising step to solve the ransomware issue. That needs to be supported by further experiments with additional samples whereby the results need to be shown in more detail.

The proposed method of Nolen Scaife et. al. (2016) proposed an early warning system for ransomware. It adds a layer between computer application and file system and analyses the operations conducted on user data what is shown in Figure 9. If it detects operations being typical for ransomware it notifies the user who can allow or decline the execution.

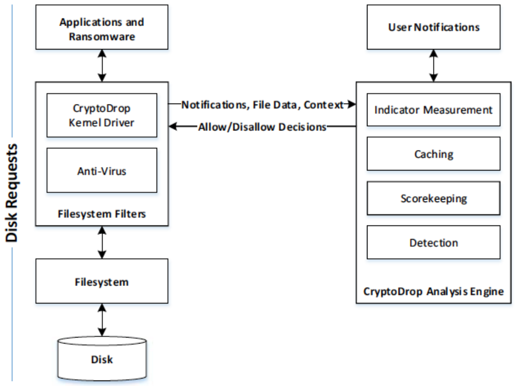


Figure 9 Architecture of the Method CryptoDrop (Nolen Scaife et. al.,2016)

In order to test this method, a virtual environment was set up including dummy user data. 2663 samples were collected belonging to 14 distinct families.

The method shows a detection rate of 100%. Zero or a few files are lost before the malware was detected.

It is claimed that the method is a very effective second line of defence when the computer system's malware detection fails. The primary limitation mentioned by the researchers is that the method cannot differentiate between normal file compression or encryption and encryption carried out by ransomware. Therefore, software tools like 7-Zip can cause false alarms.

"The conducted research was the most comprehensive study to date" (Nolen Scaife et. al. ,2016) with 14 distinct families and 2663 samples which makes the made claims of the authors credible.

Another method similarly trying to detect ransomware is proposed by Toshiki Honda et. al. (2018) since their method analyses commonly to the method of Nolen Scaife et. al. ransomware behaviour in users' real-time environments. The method creates a white list for users document editing forbidding the encryption process of ransomware on user documents. In contrast to the method of Nolen Scaife et. al. (2016), this solution does not create false alarms but it only detects two of four ransomware families before data was encrypted.

An additional massive restrict of the solution of Toshiki Honda et. al (2018) is the fact that it is only applicable for .txt files since modern ransomware is able to encrypt up to 500 different file types (Tianliang Lu et. al., 2017).

3 Comparison of Static and Dynamic Methods

The following section presents a comparison of dynamic and static solutions. The conclusions are based on the stated results but it needs to be taken into consideration that different datasets were used to test their performance. It is assumed that the results wouldn't change significantly if additional data was used.

Based on the experimental results of the evaluated proposals it is not possible to state whether static or dynamic methods detect ransomware more accurately. The method of Krzysztof Cabaj et. al. (2017) using the size of HTTP packages to detect ransomware shows a significantly higher accuracy than the dynamic method of Toshiki Honda et. al (2018). The assumption that dynamic methods are better in detecting novel variants of ransomware can also not be proven as the method of May Medhat et. al. (2018) was also able to detect ransomware samples which weren't included in its training set.

The fact that dynamic methods detect novel variants of ransomware more reliable is based on the logical fact that attackers can't alter the behaviour of ransomware as easy as its characteristics. Ransomware always needs to execute particular read and write operations in order to encrypt user data. The HTTP packages used by Krzysztof Cabaj et. al. (2017) can be changed whereas write and read operations being used by the proposal Yuki Takeuchi et. al. (2018) stay quite similar for all ransomware families.

Good results of static methods like presented from Daniel Morato et. al. (2018) base on the fact that features are found which the majority of ransomware families share but that isn't necessarily a long-term solution for the ransomware problem.

A future problem could be that ransomware could notice being executed in a sandbox environment and therefore, change its behaviour for a short period of time. That would prevent the majority of dynamic methods from detecting it. This problem was mentioned by Tianliang Lu et. al. (2017) but it has been dealt with by none of the proposals. It remains unsolved and therefore, more research is needed to improve ransomware detection methods against this probable issue in advance.

4 Recommendations

Ransomware can be countered before and after it has intruded into a computer system. Consequently, a cyber defence consisting of two dynamic methods which counter ransomware each at one stage is advised.

Therefore, a multiple line of defence consisting of the solution of Yuki Takeuchi et. al. (2018) and Nolen Scaife et. al. (2016) is advised because of their high accuracy of above 97% and their proven ability to detect novel variants of ransomware .

Thereby, ransomware is countered before an infection and additionally if the first line of defence fails. Whether these methods work together without any problems needs to be tested in further experiments.

5 Conclusions

This paper presents a critical evaluation of static and dynamic methods aimed to detect ransomware. The critical analysis of experiments and conclusions assess their validity and performance in terms of accuracy and sustainability.

The comparison of static and dynamic methods indicates that dynamic methods are the more sustainable solution due to their ability to detect novel variants of ransomware without showing worse values regarding accuracy.

The analysis shows that an accuracy of above 95% and a false positive rate below 5% are a realistic performance for individual detection methods. The elaborated knowledge leads to the recommendation of a real-world cyber defence.

Additionally, this paper shows why future research needs to be carried out to tackle the possibility that ransomware could change its behavior in sandbox-environments.

References

Aviad Cohen; Nir Nissim, 2018, ‘Trusted detection of ransomware in a private cloud using machine learning methods leveraging meta-features from volatile memory’, *Expert Systems With Applications,* Vol(102), Pages 158 – 178

Bander Ali Saleh Al-rimy;Mohd Aizaini Maarof; Syed Zainudeen Mohd Shaid, 2018, ‘Ransomware threat success factors, taxonomy, and countermeasures: A survey and research directions’, *Computers & Security,* Vol(74), Pages 144 – 166

Daniel Morato; Eduardo Berrueta; Eduardo Maga; Mikel Izal, 2018, ‘Ransomware early detection by the analysis of file sharing traffic’, *Journal of Network and Computer Application,* Vol(124), Pages 14-32

Iram Bibi1; Adnan Akhunzada; Jahanzaib Malik; Ghufran Ahmed; Mohsin Raza, 2019, ‘An Effective Android Ransomware Detection Through Multi-Factor Feature Filtration and Recurrent Neural Network’, *UK/China Emerging Technologies (UCET),* Pages 1-4

Khaled Alrawashde; Carla Purdy, 2018, ‘Ransomware Detection Using Limited Precision Deep Learning Structure in FPGA’, *IEEE National Aerospace and Electronics Conference,* Pages 152-157

Krystof Cabaj; Marcin Gregorczyk; Wojciech Mazurczak, 2018, ‘Softwaredefined networking-based crypto ransomware detection using HTTP traffic characteristics’, *Computers and Electrical Engineering,* Vol(66), Pages 353-368

May Medhat; Samir Gaber; Nashwa Abdelbaki; 2018, ‘A New Static-based Framwork for Ransomware Detection’, *IEEE 16th Int. Conf. on Dependable, Autonomic & Secure Comp.,* Pages 710-715

Nolen Scaife; Henry Carter; Patrick Traynor; Kevin R.B. Butler, 2016, ‘CryptoLock (and Drop It): Stopping Ransomware Attacks on User Data’, *IEEE 36th International Conference on Distributed Computing Systems,* Pages 303-312

Tianliang Lu; Lu Zhang; Shunye Wang; Qi Gong, 2017, ‘Ransomware Detection Based on

V-detector Negative Selection Algorithm', *International Conference on Security, Pattern Analysis and Cybernetics (SPAC)*, Pages 531-536

Toshiki Honda; Kohei Mukaiyama; Takeharu Shirai; Tetsushi Ohki; Masakatsu Nishigaki; 2018, 'Ransomware Detection Considering User's Document Editing', *IEEE 32nd International Conference on Advanced Information Networking and Applications*, Pages 907-914

Yuki Takeuchi; Kazuya Sakai; Satoshi Fukumoto, 2018, 'Detecting Ransomware using Support Vector Machines'*, ICPP '18 Proceedings of the 47th International Conference on Parallel Processing Companion,* Pages 1-6

An Evaluation of Defence Methods against HTTP DoS Attacks

Hieu Nguyen

Abstract

Denial-of-service (DoS) attacks are becoming more prevalent than ever, prompting several studies to propose and develop methods of defending against said attacks. This paper compares and evaluates several of those methods, with a focus on defence mechanisms against HTTP Flood Attacks and Slow Rate HTTP DoS Attacks. A comparison of these mechanisms reveals several limitations, wherefore future research can seek to develop further or improve the detection and mitigation methods against HTTP DoS Attacks.

1 Introduction

DoS attacks are a worldwide phenomenon. The total number of attacks is predicted to reach 17 million by 2020 (Hostingtribunal.Com, 2019). In 2018, the financial implication of these attacks reportedly amounted to $221,836.80 per hour of downtime on average (Netscout, 2018). Meanwhile, Great Britain has reported being on the receiving end of up to 1.20% of the global number of attacks in the second quarter of 2019, which increased to 1.74% in the third quarter of the same year (Karspersky, 2019).

HTTP-based attacks were the third most common type of all DoS attacks and have been for quite some time. While there are practical solutions to safeguard against these HTTP DoS attacks (Hirakawa et. al. 2016), combating them remains challenging. With the rising of defence mechanisms emerge other more complicated means of attacking, alongside the traditional HTTP Flooding Attacks. Several such means include Slow Rate (Tripathi and Hubballi, 2018) and SlowDrop (Cambiaso et. al. 2019). Consequently, there is an obvious need for continuous improvements of existing defence mechanisms as well as the developments of more innovative ways to combat such attacks.

Over the years, various methods have been proposed to recognise and protect victim servers against HTTP DoS attacks at the Application layer of Open Systems Interconnection model (OSI model) (Jaafar et. al. 2019) as well as other lower layers.

In this research paper, we aim to put forth recommendations on strengthening network security and improving the defence mechanisms against malicious traffic, by comparing and analysing the current working solutions designed for detecting and defending against HTTP DoS attacks. Earlier flood-based approaches, as well as more recent techniques such as Slow Rate will be explored.

This paper is organised as follows. Section 2 discusses the current systems of detection and mitigation against each type of HTTP DoS attack, furthered by critical analysis of the studies and comparison of the methodologies with which they were established. Conclusions and recommendations for future studies can be found in section 3.

2 Evaluation of proposed and working defence systems

This section discusses various defence systems that have been either proposed or developed and deployed. These systems will be categorised into two different types of HTTP DoS attacks: HTTP Flood and Slow Rate attacks.

2.1 Defence systems against HTTP Flood Attacks

An HTTP GET flood attack happens at the Application level, largely aiming to mimic traffic from a human user. HTTP flooding attacks are characterised under three categories: Session flooding, Request flooding and Asymmetric attack.

A method of mitigation against HTTP Flood attacks is proposed by Watanabe et. al. (2015), aiming to shift the focus on the server resources, instead of on the CPU and the memory. This proposed model grew out of the assumption that there might not be a correlation between CPU resource and the request error rate. It selected to either control the CPU or the memory resources in advance and sought to reach a pre-determined request error rate despite the varying amount of F5-attacks.

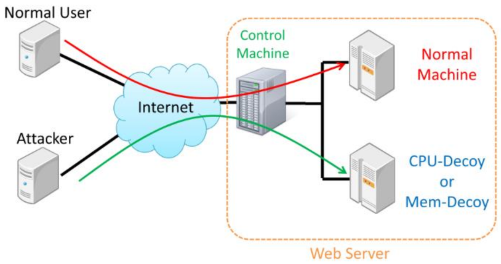


Figure 1. Illustration of the model proposed by Watanabe et. al. (2015)

Virtual hosting was employed to host multiple websites, each of which consisted of a Normal and a Decoy Machine. The model was simulated using virtual machines for the Control Machine, Normal and Decoy Machines. The system operation incorporated two modes: a normal mode and a prevention mode. Based on the NAP mappings registered by the attacker judgment program, the system will activate the prevention mode, should a user be identified as an attacker. Their requests were then sent to the Decoy Machine.

Overall, the experiment was conducted based on the authors' claim that the proposed method was effective in stably controlling the request error rate, regardless of the type of resources, successfully masking the attacker's failure. The results validated said claim as the newly founded method allowed for control of different types of resources. Notwithstanding, Watanabe et. al. (2015) pointed out that only two types of resources were tested in the experiment: CPU and memory resources. As such, there remain several other types of resources to be examined by future research, for which this study provides a strong foundation.

On the premise of the aforementioned research by Watanabe et. al. (2015), Kobayashi et. al. (2016) developed a defence mechanism such that it can go undetected by the attacker. This method aims to prevent the attacker from changing their tactic, ensuring that legitimate users aren't affected. It was tested using a network consisting of a Control Machine, Normal Machines and Decoy Machines.

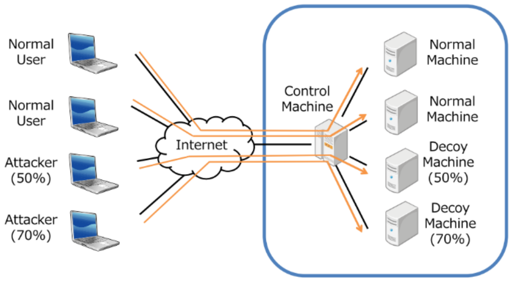


Figure 2. Illustration of the model proposed by Kobayashi et. al. (2016)

The experiment was carried out using virtual machines, given the high costs of physical machines. A host machine, which acted as the Control Machine, was used to set up four virtual machines, which acted as the Normal and Decoy Machines. As a client's requests reached the Control Machine, they were forwarded to either a Normal or a Decoy Machine. Beside its load balancing mechanism, the Control Machine also includes a Main Controller, which served to detect and mitigate against attacks.

The results from the experiment were in line with the authors' claim that the proposed method worked effectively in masking its own effectiveness in detecting the attack, leading the attacker to believe that they had succeeded. These outcomes also verified the proposed method's ability to keep genuine users from being affected by the attacks. However, as pointed out by Kobayashi et. al. (2016), the experiment was built on the assumption of a particular attack scenario, where the attackers monitored the victim server and controlled the attack rate accordingly. As such, there remain several other scenarios where the attacks are highly disruptive and constant, against which the proposed method might be incapable of mitigating.

Within each experiment discussed above, the authors conducted their testing using virtual machines where either legitimate or malicious traffic is simulated. It is reported that such virtual machines are often more susceptible to DoS attacks and the impact of said attacks on their performance is likely to be intensified, causing a higher rate of degradation than their physical counterparts (Shea and Liu, 2012). Watanabe et. al. (2015) and Kobayashi et. al. (2016) employed virtualisation software such as KVM and Xen to implement one or multiple units of virtual machines for testing. This potentially means that the virtual environments could not emulate the real system properly and that the results generated may not be so conclusive. Given the high costs of physical machines, virtual machines are likely to remain a more popular choice among researchers. Modifications are thus necessary to help create virtual environments that match their real counterparts.

2.2 Defence systems against HTTP Slow Rate DoS Attacks

Slow DoS attacks were invented as a way to circumvent current Intrusion Detection Systems and to reduce the usage of bandwidth (Cambiaso et. al. 2012), making it a more dangerous and versatile method of attack. Due to the reduced usage of resources, carrying out a Slow DoS attack is easier than ever, made possible even on low-performance hosts such as routers and mobile devices (Cambiaso et. al. 2019). One type of Slow HTTP DoS Attacks emerging over the past few years is Slow Rate.

A Slow Rate DoS attack is characterised by the releases of incomplete requests by the attacker. Due to the slow rate that incomplete requests usually interact with the server, they get stored in the queue, holding back legitimate requests from realising themselves (Tripathi and Hubballi, 2018). Slow Rate attacks are also categorised under three main types: Slow Read, Slow Header and Slow Body.

Slow Read DoS Attacks

On the one hand, a simple Slow Read DoS Attack can be prevented using ModSecurity (Park et. al. 2015) or other configurations, built-in or otherwise, specific to each web server application (Suroto, 2017). It’s worth noting, however, that these methods can be less successful in preventing Distributed DoS Attacks (DDoS). The prevention of HTTP DDoS Attacks is more effective by ‘disconnecting the attack connections selectively by focusing on the number of connections for each IP address and the duration time’ (Hirakawa et. al. 2016).

On the other hand, one of the earliest methods of detecting Slow HTTP Dos Attacks was proposed by Duravkin et. al. (2014), based on the premise that most current mechanisms don’t protect the servers in their entirety. They put forward a model which consisted of 6 modules: (1) collecting traffic, (2) calculating various traffic parameters, (3) generating network statistics, (4) calculating web server parameters, (5) marking traffic potentially from an attacker and (6) classifying the attack. Duravkin et. al. (2014) claimed that this proposed method allowed for detecting the presence of malicious traffic as well as identifying where it originates.

This study differed to most others analysed in this paper in that the expected result was a model, which included several modules and, theoretically, was built to systematically detect potential attacks and identify their sources. Experiments were yet to be conducted in order to test and confirm the extent to which the proposed model could deliver the desired outcomes. Notwithstanding, this study provided a strong momentum for future research to affirm the utility of the model and expand its scale.

Another countermeasure for Slow HTTP DoS Attacks involves putting in place a maximum request duration time, after which the request should be automatically terminated, freeing up server resources as a result (Hong et. al. 2017). This proposed method grew out of the need to address the limitations of earlier host-based mitigation mechanisms, which required access to the server under attack. The proposed framework consisted of 3 phases: (1) detection phase, where the system works to find the attacker; (2) identification phase, where the monitor identifies the attacker and passes the information on to the SDN controller; and (3) mitigation phase, where the SDN controller sends a request to the switches to block further attempts to connect by the attacker and at the same time, terminating their connections.

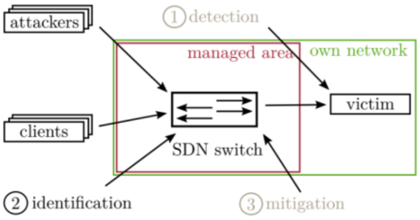


Figure 3. The framework architecture (Hong et. al. 2017)

Several schemes were implemented as part of the framework, each configured with specific parameters. Three data sets were evaluated. One was a training data set, used to determine the appropriate threshold for each scheme. The second data set was used to indicate whether the trained system could mitigate against the attack adequately. The last data set was used to indicate whether the parameters could be deployed universally.

The results justified the authors' claim that while the schemes didn't exhibit a level of accuracy high enough to leave the system active all the time, they were an effective reactive defence system against ongoing Slow attacks, nonetheless.

Slow Header and Slow Message Body DoS Attacks

Another detection system was proposed by Tripathi et. al. (2016). It employs an anomaly-based technique that quantifies the similarity between two profiles, or probability distributions, created in a training phase and testing phase, using Hellinger distance metric. In the training phase, the test generated a probability distribution from normal HTTP traffic that consisted of both complete and incomplete requests. In the testing phase, said profile was compared with other probability distributions generated from Slow Header and Slow Message Body attack intervals of the traffic, respectively. Tripathi et. al. (2016) then confirmed that the proposed method was able to detect said attacks with a high level of accuracy. Two experiments were run, one with simulated HTTP traffic and the other with real traffic from the authors' institution's public web server.

The experimental results were in agreement with the authors' claim as they showed a significant spike in incomplete GET requests during a Slow Header Attack and in incomplete POST requests during a Slow Message Body Attack. Overall, based on the favourable outcomes of the experiments, not only was the method proposed by Tripathi et. al. (2016) capable of detecting Slow DoS Attacks with a high level of accuracy, but also difficult to evade. As stated by the authors, evading was only possible if the distribution profile of the attackers' HTTP requests was similar to the one derived from the training period.

The methods of detection proposed by Tripathi et. al. (2016) utilised an anomaly-based technique, which was able to detect new attacks without the training required in misuse-based detection systems. However, Sekar et. al. (2002) warn that anomaly-based detection is often susceptible to a high degree of false alarms, which can be minimised if used in conjunction with specification-based techniques. A specification-based detection method, though just as effective in detecting new attacks, can require an extensive number of manual configurations to help capture the behaviours of legitimate users and differentiate them from those of an attacker. The detection method proposed by Duravkin et. al. (2014) employed such a technique, capable of detecting Slow attacks with a low degree of false alarms.

3 Conclusions

In response to the increasing number of HTTP DoS attacks worldwide, numerous studies have put forward proposals of mechanisms that could potentially detect these attacks and help protect the victim servers. In analysing and comparing these methods, we find that they provide a strong foundation for the developments of measures that can effectively detect malicious traffic and improve the overall network security. However, there are several limitations still to be addressed before a defence mechanism can be entirely relied on.

One limitation involves the detection techniques, whereby a model can work effectively autonomously or in conjunction with another. Future studies should aim to experiment with various combinations of techniques. In so doing, they can maximise the full potential of existing research and uncover new ways to combat DoS attacks without reinventing the wheel.

Another limitation has to do with the environments in which these models are tested; whether they correspond with the highly complex environments in the real world. Future research can look at improving the testing environments with modifications such that they simulate a real-world environment and consequently, deliver results with better accuracy and higher reliability.

References

Cambiaso, E., Papaleo, G. and Aiello, M., 2012, ‘Taxonomy of slow DoS attacks to web applications’, *International Conference on Security in Computer Networks and Distributed Systems*, Springer, pages 195-204.

Duravkin, I., Loktionova, A. and Carlsson, A., 2014, ‘Method of slow-attack detection’, *First International Scientific-Practical Conference Problems of Infocommunications Science and Technology*, IEEE, pages 171-172.

Hirakawa, T., Ogura, K., Bista, B. B. and Takata, T., 2016, ‘A defense method against distributed slow HTTP DoS attack’, *19th International Conference on Network-Based Information Systems (NBiS)*, IEEE, pages 152-158.

Hong, K., Kim, Y., Choi, H. and Park, J., 2017, ‘SDN-assisted slow HTTP DDoS attack defense method’, *IEEE Communications Letters*, 22**,** pages 688-691.

Hostingtribunal.Com (2019) *39 Jaw-Dropping DDoS Statistics to Keep in Mind for 2019.* Available at: https://hostingtribunal.com/blog/ddos-statistics/ (Accessed: 21 November 2019).

Jaafar, G. A., Abdullah, S. M. and Ismail, S., 2019, ‘Review of Recent Detection Methods for HTTP DDoS Attack’, *Journal of Computer Networks and Communications*, pages 1-10.

Karspersky (2019) *DDoS attacks in Q3 2019.* Available at: https://securelist.com/ddos-report-q3-2019/94958/ (Accessed: 21 November 2019).

Kobayashi, R., Otani, G., Yoshida, T. and Kato, M., 2016, ‘Defense method of HTTP GET flood attack by adaptively controlling server resources depending on different attack intensity’, *Journal of Information Processing*, 24**,** pages 802-815.

Netscout (2018) *Cloud in The Crosshairs*. Available at: https://www.netscout.com/report/ (Accessed: 21 November 2019).

Park, J., Iwai, K., Tanaka, H. and Kurokawa, T., 2015, ‘Analysis of slow read dos attack and countermeasures on web servers’, *International Journal of Cyber-Security and Digital Forensics (IJCSDF)*, 4**,** pages 339-353.

Sekar, R., Gupta, A., Frullo, J., Shanbhag, T., Tiwari, A., Yang, H. and Zhou, S., 2002, ‘Specification-based anomaly detection: a new approach for detecting network intrusions’, *Proceedings of the 9th ACM conference on Computer and communications security*, ACM, pages 265-274.

Shea, R. and Liu, J., 2012, ‘Understanding the impact of denial of service attacks on virtual machines’, *Proceedings of the 2012 IEEE 20th International Workshop on Quality of Service*, IEEE Press, 27, pages 1-9.

Suroto, S., 2017, ‘A review of defense against slow HTTP attack’, *JOIV: International Journal on Informatics Visualization*, 1**,** pages 127-134.

Tripathi, N. and Hubballi, N., 2018, ‘Slow rate denial of service attacks against HTTP/2 and detection’, *Computers and security*, 72**,** pages 255-272.

Tripathi, N., Hubballi, N. and Singh, Y., 2016, ‘How secure are web servers? An empirical study of slow HTTP DoS attacks and detection’, *11th International Conference on Availability, Reliability and Security (ARES)*, IEEE, pages 454-463.

Watanabe, M., Kobayashi, R. and Kato, M., 2015, ‘HTTP-GET Flood Prevention Method by Dynamically Controlling Multiple Types of Virtual Machine Resources’, *Journal of Information Processing*, 23**,** pages 655-663.

A Critical Evaluation of Current Techniques Aimed at Improving Physical Layer Security in Wireless Network

Lip Chuan Sui

Abstract

Wireless Network is a growing area in modern technology. However, the security of wireless network faces threats because of wireless broadcasting nature. Therefore, Physical Layer Security (PLS) is a promising approach in collecting data with secure and safe. This paper provides analysis of some of researches conducted for improving physical layer security, which are relay selection, Artificial Noise (AN) and Cooperative jamming. Each method is critically analyzed to discover strengths, weakness and validity. Comparison of methods is carried out to determine which method is promising and suggestions are made for further research on improving each method.

1 Introduction

Wireless Networks such as Mobile Network, Internet of Things (IoT) are influencing our lifestyle from how we react to the way we behave. Wireless communication faces serious threat from eavesdropping attacks by unauthorized things (eavesdroppers) because of open nature of wireless medium. Therefore, it is essential to secure the legitimate transmission

Almalkawi et al. (2019) proposed a lightweight and efficient encryption scheme based on chaotic algorithms to efficiently encrypt digital images. Shukla (2017) proposed a machine learning algorithm which would detect wormhole attacks in Internet of Things.

Zhang 2016 states that physical-layer security technology could serve as a promising complement to the encryption protocols to enhance the security performance of wireless communications.

In recent years, physical layer security (PLS) has shown great potential in securing safety in communication if wireless network. Exploiting the inherent randomness and difference of wireless channels, eavesdroppers could be kept at bay regardless of their computing capabilities.

The aim of this research paper is to critically evaluate and analyze the current research that has been done in order to improve security of wireless network through physical layer. This paper focus on three methods which are relay and jammer selection, injection of noise or jamming signal and unmanned aerial vehicle. These methods will be analyzed thoroughly in order to gauge their effectiveness, and conclude with some recommendations on how they can be improved.

2 Current Techniques used in Physical Layer Security

This section reviews relay and jammer selection, injection of noise or jamming signal and unmanned aerial vehicle researches carried out by different researchers. The method and the validity of the experiments as well as the implications of the results will be discussed in this paper.

2.1 Relay and Jammer Selection Technique

Let’s assume there are usually 3 nodes in a simple wireless network, which are a source, a destination and an eavesdropper. Relay selection Method provides a helper node to improve security by either acting as information relay or signal jammer.

Van et al. (2018) proposes a method to transmit data signal securely with the help of multiple intermediate nodes. The intermediate nodes act as either conventional relays or as jammers to degrade the channel of eavesdroppers. Van et al. (2018) also implements energy harvest (EH), which enable sensor nodes harvest energy from multiple power transfer stations (PTSs). The source uses this harvested energy to transmit information to the base station (BS) continuously. Meanwhile, one of the relays uses harvested energy to emit jamming signal to confuse the eavesdroppers.

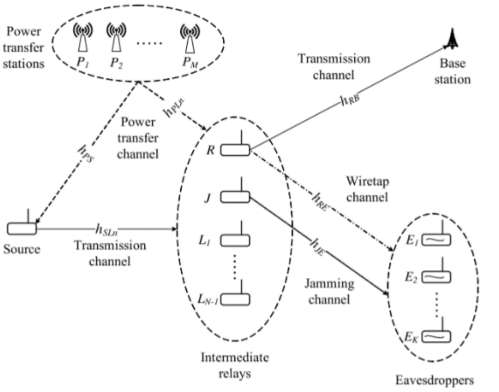


Fig 1 Wireless network with Power transfer stations (Van et al. 2018).

Van et al. (2018) derives an exact closed-form expression for the Secrecy Outrage Possibility (SOP). They considered many factors that would affect SOP They compared three schemes which are Best-relay-and-random-jammer scheme (BRS), Random-relay-and-best-jammer scheme (RBS) and Best-relay-and-best-jammer scheme (BBS). They run multiple simulations using Monte Carlo to observe how SOP changes with other variables.

Their numerical results indicate that the BBS outperforms both the BRS and the RBS. Best secrecy performance could be achieved by choosing best relay and best jammer in a scheme. They concluded that the performance of the proposed scheme improves as the number of relays and PTS increase.

Simulations run by Van et al. (2018) did not include scheme without PTS. Comparisons only be made on schemes with different selections criteria without including energy harvest. Therefore, the research paper did not show enough evidence to proof the benefit or differences of implementing EH technique. However, the research was conducted on multiple variables without any bias towards the proposed method. They also mention what software used for simulation. Therefore, they have enough evidence to proof their method is valid.

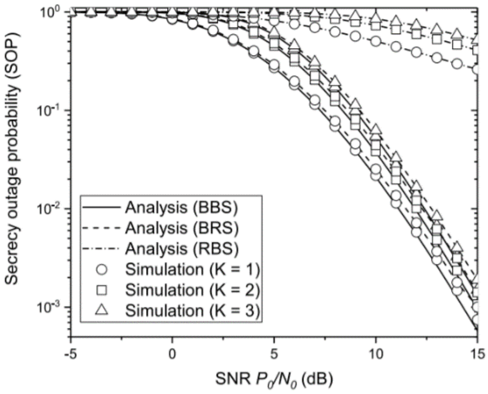


Fig 2 Result shows that BBS has lowest SOP (Van et al. 2018).

Wang et al. (2016) considered the aspect of untrusted relay nodes. Besides forwarding the received signal, untrusted nodes may also decode signal without permission. They proposed to use the successive relaying scheme to secure an amplify-and-forward relay network with multiple untrusted nodes. There is no direct link between Source and Destination. Two relays are chosen to transmit data signal. The inter-relay interference (IRI) is used to cause interference to untrusted relays.

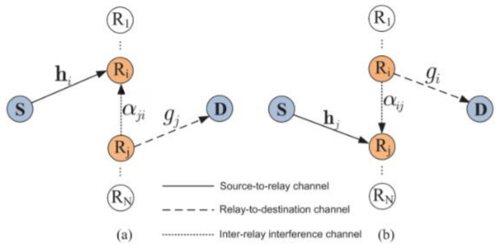


Fig 3 Successive relaying scheme to counter untrusted relay (Wang et al. 2016).

They assume that relay nodes are both helpers and eavesdroppers. Wang et al. (2016) presented different relay selection schemes with each focus on specific criteria. Simulations considered many factors including Rayleigh fading, number of antennas and others. All the factors are remained same for all schemes during simulation test to prevent bias.

Simulation results indicates that all successive relaying schemes outperform the conventional single-path (SP) relaying scheme. Results further indicate that it is favorable to increase the number of untrusted relay nodes to further improve the results. Therefore, their scheme is better than conventional scheme.

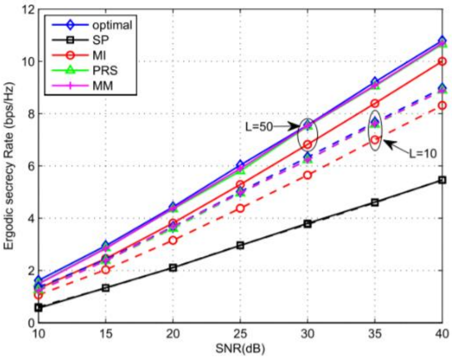


Fig 4 Result indicates that all the three different successive relaying schemes achieves higher secrecy rates than conventional single-path (SP) relaying scheme. (**Wang et al. 2016).**

Wang et al. (2016) did not include any simulation software, algorithm or methods to run simulation in the paper but they did match the result with Monte Carlo simulations. However, multiple variables were considered when conducting simulations to prevent any bias. Both single path and successive relay selection are simulated to show clear difference. Therefore, the results are justified.

In conclusion, further research should be done by combining two methods discussed above. Van et al. (2018)'s research could solve energy problems and be effective in wireless medium where devices are resource constrained. Wang et al. (2016) considered the aspects of untrusted relay, which is different from any other researches about relay selection methods. This method could be implemented in wireless network where relays are easily compromised by hacker. Relay and jammer selection method could be further improved by improving both of these techniques.

2.2 Injection of Noise or Jamming Signal

The injection of artificial noise (AN) or jamming signal method is to emit AN or jamming signal to confuse the eavesdropper.

In Fig 5, Hu et al. (2018) considered an IoT network consists of a controller, an actuator, a cooperative jammer and multiple eavesdroppers. An actuator is device used to make changes in physical world. The cooperative jammer generates jamming signals to confuse eavesdroppers.

They proofed their scheme by using mathematical formula without mentioning any simulation software or algorithm. They run a controlled experiment with and without CJ. They develop lots of formulas to derive SOP and power allocation ratio to test the performance with and without CJ.

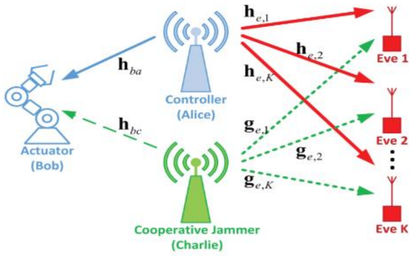


Fig 5 Wireless network model of controller and actuater (Hu et al. 2018).

Their mathematical results indicate that SOP in CJ scheme is lower than that without CJ. They observed that jammer causes significant performance degradation to eavesdroppers. Therefore, they concluded that the SOP performance is enhanced with CJ scheme.

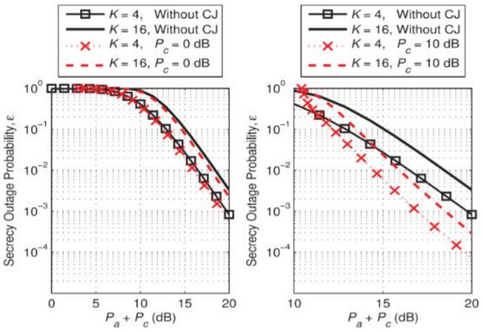


Fig 6 Result show that SOP is lower with CJ (Hu et al. 2018).

This study conducted by Hu et al. (2018) did not use any simulation software or algorithm but only mathematical formula to calculate the results. They explained clearly their formulas but the study is still lacking validly. Thus, there is not enough evidence to proof their results valid.

Zhang et al. (2018) implemented energy harvest technique into cooperative jammer in orthogonal-frequency-division-multiplexing (OFDM) system. The cooperative jammer works by receiving harvest energy from source and using it to emit jamming signal.

Zhang et al. (2018) considered two types of receivers at the destination, namely Type-I and Type-II receivers, with and without the capability of canceling the jamming signals from the jammer, respectively. They maximize the secrecy rate via joint time and power allocation. They compared the results from these two types of receivers.

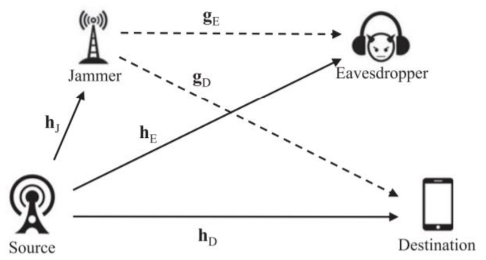


Fig 7 Wireless Network Model with Source, Destination, Jammer and Eavvesdroppers (Zhang et al. 2018).

Simulation results show that the proposed approaches achieve higher secrecy rate than conventional schemes. They observed all schemes with CJ outperform the conventional scheme. Therefore, they concluded that their methods is very effective in improving PLS.

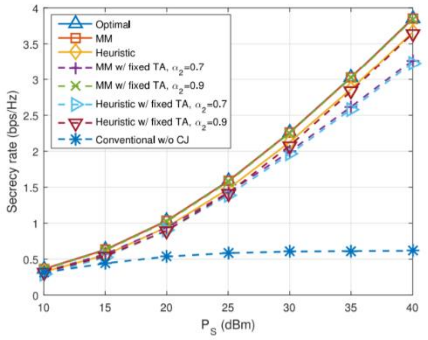


Fig 8 Result shows that all schemes with jamming achieve higher secrecy rate than the conventional scheme. (Zhang et al. 2018).

The research conducted by Zhang et al. (2018) considered the security of receiver by using two different types of receiver with and without cancelling jamming signal ability. However, the simulations only compare the results of Type-I and Type-II receivers. There is not enough evidence to show clear advantages of their methods. They didn't explain how they obtain simulation of conventional scheme without cooperative jamming. They didn't mention any simulation software or algorithm. Therefore, their results are not valid.

In conclusion, injection of AN or jamming signal is an advanced technique which requires further improvement because it will increase computational complexity. It is favorable to include energy harvest technique because the technique consumes power. Besides, it is favorable to use CJ technique because it eliminates the needs of source to emit AN or jamming signal, which directly decrease throughput of data signal. This technique works best if it combines with other technique such as relay selection or UAV. However, more time is needed for designing suitable jamming signal which will not affect original data signal. Besides, Hu et al. (2018)'s paper studied controller and actuator, which is quite rare mentioned in IoT wireless network. It is important to look into other aspect of wireless network, rather than focus solely on source to destination.

2.3 Unmanned Aerial Vehicle

Unmanned aerial vehicles (UAVs) provides many advantages such as wide coverage and high mobility. Hence, it is practical to apply UAV in wireless network.

Zhang et al. (2019) exploit the high mobility of UAVs to establish stronger links with legitimate ground nodes via joint trajectory and power control optimization.

Zhang et. al. (2019) considered a simplified three-node UAV-ground communication system as shown in Fig. 9, where a UAV at fixed altitude communicates with a ground node in the present of an eavesdropper. They considered both UAV-to-ground (U2G) and ground-to-UAV (G2U) links. They run simulation to compare four different benchmark schemes which are two types of trajectory optimization with and without power control algorithm.

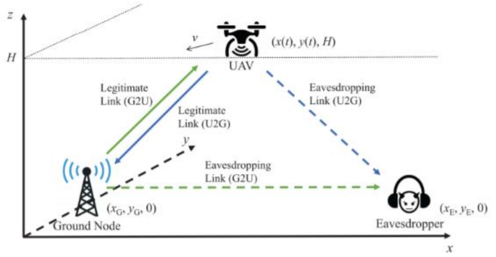


Fig 9 Wireless network consisting of a UAV, a ground node and a potential eavesdropper (Zhang et al. 2019).

Their results indicates scheme with default trajectory optimization with power control achieves highest secrecy rates. They concluded that UAV should adapt different algorithms based on circumferences to achieve efficiency. Besides, in the high transmit power regime, trajectory optimization is more significant while in the low transmit power regime, power control is more important.

This study conducted by Zhang et al. (2018) did not mention any simulation software or algorithm to produce their results. However, they displayed clear results of simulation between four different types of schemes. They further compared the results to show differences with or without their method. Subsequently, they could conclude that their schemes could be applied depending on different circumstances for efficiently. Hence, they have enough evidence to proof their results valid.

Shang et. al. (2019) proposed a method to apply UAV into vehicle-to-everything (V2X) services. UAV could act as trusted relays to provide Line-of-Sight (LoS) transmissions. Furthermore, UAV can use camera and radar to detect potential vehicle eavesdroppers. UAV could act as friendly jammer.

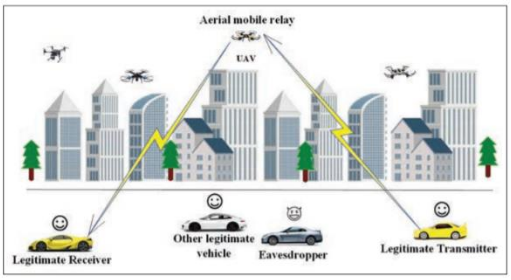


Fig 10 UAV in V2V communications (Shang et al. 2019).

Shang et. al. (2019) compared vehicle-to-vehicle (V2V) communication under three cases which are with UAV jammer, with ground jammer and without jammer. They run simulation twice, one with constant velocity of flight of UAV while one with increasing velocity. They assume that the velocities of all nodes are constant.

Results indicated that application of UAV jammer has the highest secrecy rates. For the Fig 10, Shang et. al. (2019) observed that the secrecy rate increases but later decreases. This is due to when UAV is slow, it is difficult to catch up with eavesdropper. When the velocity increases, UAV may pass eavesdropper quickly and decrease the time duration of V2V communication, thus decrease secrecy rates. They colluded that their method could improve secrecy performance.

This study conducted by Shang et al. (2018) did not mention any simulation software to produce their results. They also did not show any calculation and formulation for their works. They only explain their methods with words without enough evidence. Hence, their results are invalid. However, they further elaborate the challenges and research opportunities for UAV which are including trajectory design, energy efficiency and others.

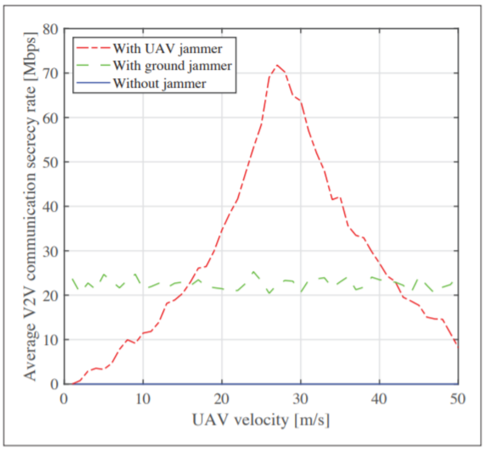


Fig 11 Results from simulation with increasing velocity of UAV (Shang et al. 2019).

In conclusion, UAV has advantages over conventional ground relays because UAV could move freely in the sky and provide larger area of trajectory. The research conducted by Zhang et al. (2019) exploit UAV mobility to optimize joint trajectory and transmit power control using only 1 UAV. The method could be further enhanced by increasing the number of UAV. For example, Wu et al. (2018) proposed a multi-UAV enabled wireless communication system. Moreover, UAV could be applied to V2X service. This show that UAV could be applied into various types of wireless network to improve PLS. This method shows a lot of promise to enchance the security of wireless network.

3 Comparison and Evaluation of Techniques in Physical Layer Security

This section presents comparison of the three methods mentioned above.

Relay and jammer selection methods rely on relays to redirect the data signal to destination and jammers to emit AN or jamming signal to degrade eavesdroppers. Compared to injection of AN and jamming signal methods, relays selection required independent devices to operate, which will cost more money, time and management. These jammers also required extra management to control to allow CJ happen. However, relays selections allow communication between source and destination which are far apart while injections of AN or jamming signal method could not.

For injection of AN and jamming signal, throughput of data signal has to be compromised to add in AN or jamming signal. Besides, Xu et al. (2016) states that AN is not suitable for IoT applications in practice because it consumes high amount of energy as well as causes interference to adjacent nodes. Hence, it is favorable to combine both relays selection and injection of AN or jamming signal to solve each problem.

UAV should be implemented with EH technique as shown by Van et al. (2018). Guo et. al. (2019) stated that EH is better than solar energy because it is less dependent on weather. UAV rans by battery and has limited time to operate. With this technique, UAV could operate longer time.

Both three ideas could be combined into one great method. Imagine relay selection and noise injection methods but the devices are flying in the sky. In Fig 11, Cai et al. (2018) applied two UAVs in this system where UAV 1 has communication with users while UAV 2 emit jamming signal to interference with eavesdroppers.

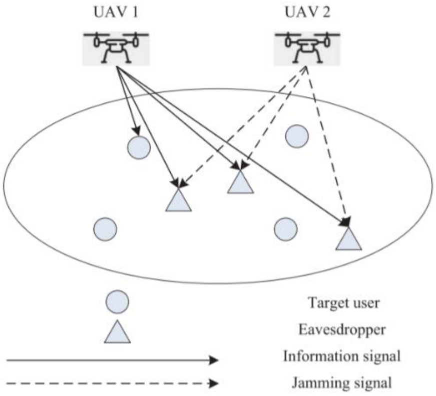


Fig 12 UAV acts as friendly jammer. (Cai et al. 2019).

It clearly shows that relay and jammer selection is the best method among three of them. Relay and jammer selection could transmit data signal more securely while implementing EH and AN. It also achieves cost, energy and throughput efficiency.

4 Conclusions

This paper evaluated different researches carried out to improve PLS in wireless network. The researches included injection of AN or jamming signal, relay and jammer selection and UAV. Upon critically evaluating the above-mentioned techniques, it clearly shows that relays and jammer selections is the best method among three of them. This is because relay and jammer selection method implement helper nodes to redirect data signal and acts as jammer. This allows the distribution of tasks among the devices and avoids Source and Destination nodes using too much computational complexity. This method also provides larger distance to data signal transmission. Besides, EH technique could improve energy efficiency. Even though relay and selection method require higher cost because of additional devices.

The results of this research will benefit the networking industrial because it could improve security of wireless communication. PLS is a great approach to improve security besides encryption or AI. This will provide secure wireless communication and data collection without the risk of eavesdropping.

One of the side effects of injection of AN or jamming signal is that the system requires high usage of energy and computational power. Thus, injection of AN or jamming signal truly shines when it is incorporated with other methods.

UAV method is promising but it stills need much more researching and monitoring. Gatwick Airport drone incident shows that 1 drone without supervision could bring devastating effect. The consequence could be devastating if UAV is not handled properly.

References

Almalkawi I. T., Halloush R., Alsarhan A., Al-Dubai A. and Al-karaki J. N., 2018, 'A lightweight and efficient digital image encryption using hybrid chaotic systems for wireless network applications', *In Journal of Information Security and Applications,* Vol 49 Dec.

Cai Y., Cui F., Shi Q., Zhao M. and Li G.Y., 2018, 'Dual-UAV-Enabled Secure Communications: Joint Trajectory Design and User Scheduling', *IEEE Journal on Selected Areas in,* Sep 36(9):1972-1985.

Guo H., Yang Z., Zou Y., Tsiftsis T., Bhatangar M.R. and Lamare R.C.D., 2019, 'Secure Beamforming for Cooperative Wireless-Powered Networks With Partial CSI', *IEEE Internet of Things Journal*, Aug, 6(4):6760-6773.

Hu L., Wen H., Wu B., Pan F, Liao R., Song H., Tang J., and Wang X., 2018, 'Cooperative Jamming for Physical Layer Security Enhancement in Internet of Things', *IEEE Internet Things Journal,* Feb, 5(1):219-228.

Nhan Vo V., Nguyen T.G., So-In, C., Baig Z.A. and Sanguanpong S., 2018, 'Secrecy Outage Performance Analysis for Energy Harvesting Sensor Networks with a Jammer Using Relay Selection Strategy', *IEEE Access,* 6:23406-23419.

Shang B., Liu L., Ma J. and Fan P., 2019, 'Unmanned Aerial Vehicle Meets Vehicle-to-Everything in Secure Communications', *IEEE Communications Magazine IEEE Commun. Magazine, IEEE.* Oct, 2019, 57(10):98-103

Shukla P., 2017, 'ML-IDS: A machine learning approach to detect wormhole attacks in Internet of Things', *Intelligent Systems Conference* Sep, pages 234-240.

Wang W., Teh K.C. and Li K.H., 2016, 'Relay Selection for Secure Successive AF Relaying Networks with Untrusted Nodes', *IEEE Transactions on Information Forensics and Security,* Nov, 11(11):2466-2476.

Wu Q., Zeng Y. and Zhang, 2017, 'Joint Trajectory and Communication Design for Multi-UAV Enabled Wireless Networks', *IEEE Transactions on Wireless Communications,* March, 17(3): 2107-2121.

Xu Q, Ren P, Song H and Du Q, 2016, 'Security Enhancement for IoT Communications Exposed to Eavesdroppers with Uncertain Locations', *IEEE Access,* pages 1290–1291.

Zhang G., Wu Q., Cui Miao. And Zhang R, 2018, 'Securing UAV Communications via Joint Trajectory and Power Control', *IEEE Transactions on Wireless Communications*, Feb, 18(2): 1376-1389.

Zhang Y, Shen Y, Wang H, Yong J and Jiang X, 2016, 'On Secure Wireless Communications for Iot Under Eavesdroppers Collusion', *IEEE Transactions on Automation Science and Engineering*. 13(3), page 1281-1293.

Zhang G., Xu J., Wu Q., Cui M., Li X. and Lin F., 2018, 'Wireless Powered Cooperative Jamming for Secure OFDM System', *IEEE Transactions on Vehicular Technology*, Feb, 67(2):1331-1346.

Critical Analysis of Current Research Aimed at Improving Detection of Phishing Attacks

Eddie Wong Shou Shie

Abstract

This paper provides a critical and objective analysis of current research into the improvement of phishing detection. As many people use the Internet to conduct daily business including monetary transactions, phishers have appeared in higher frequency and with better phishing attacks to target them, making them key considerations in Internet Security. The paper will cover and evaluate Classification Model, Deep Learning and Support Vector Machine in detecting phishing attacks. Analysis shows Deep Learning techniques is highly effective at detecting phish but may suffer when taking external factors into consideration. Comparison of the various methods are carried out to determine the best performance and recommendations are made for further improvement.

1 Introduction

Phishing attacks are an increasingly important priority to factor when implementing Internet Security, as more users rely on the Internet for daily tasks which require the giving of personal information. As phishing attacks become more varied, new phishing detection techniques are needed to keep up

Mahdieh Zabihimayvan, Derek Doran(2019) proposed a detection model using Fuzzy Rough Set theory for selecting features from suspicious sites to determine legitimacy.

In a study to determine the effectiveness of similarity score in determining if a site is phish, Routhu Srinivasa Rao, Alwyn Roshan Pais(2019a) observed that phishing sites hosted on compromised servers have many differences between them and actual legitimate sites.

Carlo Marcelo Revoredo daSilva et. al.(2019) looked into the commonality of features shared by phishing attacks to better phishing predictions. They note that some features appear regularly in phishing attacks, and speculate that further investigations could lead to better phishing detection.

Tommy Chin et. al.(2018) published a method using deep-packet inspection leveraged with software-defined networking to identify and contain phishing attacks through web-based communication.

The paper aims to provide an objective and critical analysis of current research done on phishing detection improvements. The paper will focus on certain fields, namely Classification Model, Deep Learning and Support Vector Machine. The paper is organized into discussion and evaluation of methods in the specified fields, comparisons and conclusions of discussed methods. Recommendations on the best method and possible external applications will be made as well.

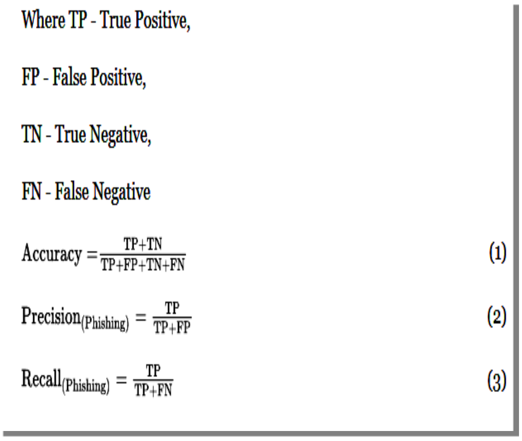
2 Phishing Attack Detection Methods

This section focuses on current proposed methods for detecting phishing attacks and an analysis of the methods. The fields analyzed are Classification Model, Deep Learning and Support Vector Machine.

2.1 Classification Model

Seow Wooi Liew et. al.(2019) presents a method that uses Random Forest classification model as a Twitter security alert mechanism. The researchers claim that the proposed model will be effective at detecting phishing threats to users in real-time.

To justify their claim, Seow Wooi Liew et. al.(2019) first tested various classification features and set a classification accuracy of 94.56% as the baseline to be achieved. Furthermore,10 fold test mode for cross validation purposes was used due to limited URLs used. Test mode usage ensured all data is utilized for training and testing.



Where TP - True Positive,

FP - False Positive,

TN - True Negative,

FN - False Negative

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

$$\text{Precision}_{(\text{Phishing})} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall}_{(\text{Phishing})} = \frac{TP}{TP+FN} \quad (3)$$

Figure 1 Standard Information Retrieval Metrics Formula, Seow Wooi Liew et. al.(2019)

The researchers then created a dataset consisting of 100 legitimate URLs from Twitter, and 100 phishing URLs from PhishTank. The gathered URLs were posted one-by-one on Twitter, and the prompting of security alerts for each URL if phish were observed and recorded. The alert is deemed accurate if the URL has the phishing label given to phishing URLs during training.

The researchers found that usage of Random Forest was effective at achieving a high classification rate at 97.50%. They also compared their results with another method that uses Random Forest. It was concluded that the proposed model has a high efficiency rate at detecting phishing URLs on Twitter.

The experiment by Seow Wooi Liew et. al.(2019) has no bias. The experiment and dataset are clearly outlined, allowing for easy reproducibility. Moreover, the results were compared with a similar model’s to ensure effectiveness. However, the dataset size could be larger, to ensure better training of the proposed model on a larger variety of data. Furthermore, no explanation was given for how the URLs from Twitter was confirmed benign, causing possibility of false positives or negatives. The proposed model may be effective in detecting phish on other social media and websites that includes links to external sites.

Ozgur Koray Sahingoz et. al.(2018) proposes an anti-phishing system that uses seven distinct classification algorithms and natural language processing features. They aim to determine which classification algorithm produces the best results when combined with natural language processing.

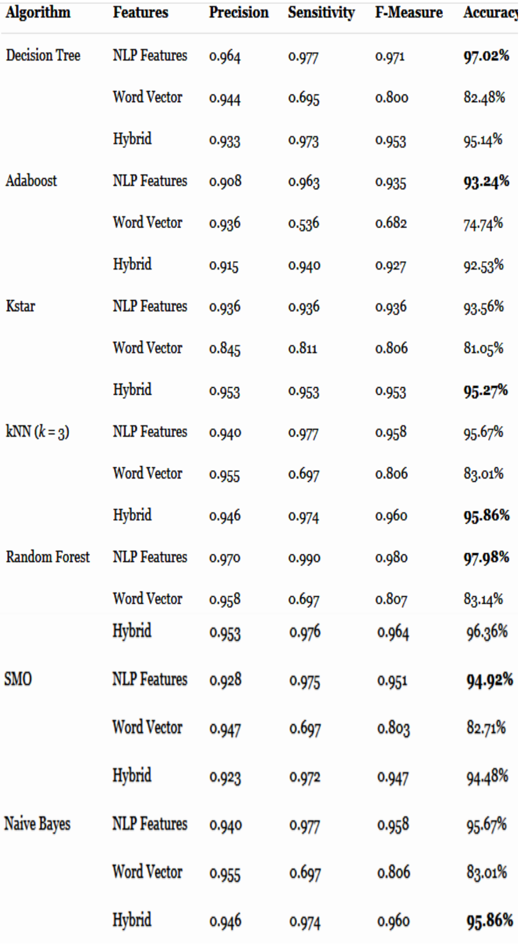
Ozgur Koray Sahingoz et. al.(2018) tested the proposed system using a dataset comprised of 73,575 URLs, approximately half being phishing URLs from PhishTank and the remainder from Yandex Search API. 10-fold cross validation in Weka was used alongside the default parameters for the algorithms during the experiment. Moreover, different features were combined with the algorithms to determine the best combination. These features are divided into three main classes as shown in Table 2.

The researchers conclude that Random Forest, combined with NLP features delivered the best results.Furthermore, the results show that word vector combinations perform poorly, and hybrids perform slightly better than NLP features in some combinations but worse in others. The researchers also note that the method is likely ineffective if the URL contains only a single domain name.

The experiments were clearly outlined with no bias towards any particular combination. The dataset used has sufficient quantity and from reputable sources. Of concern is no comparison with similar models or research was done, making the results consistency with existing statistics uncertain.The findings can be applied as well for research or experiments that requires classification models.

While all classification algorithms achieved high accuracy, Random Forest outperforms the others, especially combined with NLP features. Furthermore, Random Forest achieves high accuracy in real life applications as shown by Seow Wooi

Liew et. al.(2019). Both papers evaluated however had incomplete testing, rendering their concluded results inconclusive.



Algorithm	Features	Precision	Sensitivity	F-Measure	Accuracy
Decision Tree	NLP Features	0.964	0.977	0.971	**97.02%**
	Word Vector	0.944	0.695	0.800	82.48%
	Hybrid	0.933	0.973	0.953	95.14%
Adaboost	NLP Features	0.908	0.963	0.935	**93.24%**
	Word Vector	0.936	0.536	0.682	74.74%
	Hybrid	0.915	0.940	0.927	92.53%
Kstar	NLP Features	0.936	0.936	0.936	93.56%
	Word Vector	0.845	0.811	0.806	81.05%
	Hybrid	0.953	0.953	0.953	**95.27%**
kNN ($k = 3$)	NLP Features	0.940	0.977	0.958	95.67%
	Word Vector	0.955	0.697	0.806	83.01%
	Hybrid	0.946	0.974	0.960	**95.86%**
Random Forest	NLP Features	0.970	0.990	0.980	**97.98%**
	Word Vector	0.958	0.697	0.807	83.14%
	Hybrid	0.953	0.976	0.964	96.36%
SMO	NLP Features	0.928	0.975	0.951	**94.92%**
	Word Vector	0.947	0.697	0.803	82.71%
	Hybrid	0.923	0.972	0.947	94.48%
Naive Bayes	NLP Features	0.940	0.977	0.958	95.67%
	Word Vector	0.955	0.697	0.806	83.01%
	Hybrid	0.946	0.974	0.960	**95.86%**

Table 1 Comparison of classification algorithms combined with features, Ozgur KoraySahingoz et. al.(2018)

2.2 Deep Learning

The advancement of deep learning techniques caused the proposal of deep learning techniques in detecting phish. Bo Wei et. al.(2019) propounds a method that uses a deep learning algorithm, emphasizing real-time detection while simultaneously being energy efficient.

The method was tested by Bo Wei et. al.(2019) using a dataset containing 1,523,966 URLs, where 999,996 were legitimate URLs taken "from the list of Alexa top 1 million sites, hphosts, Joewein, malwaredomains, and phishtank"(Bo Wei et. al., 2019) and the remainder being phishing URLs. Repeat URLs are removed from the dataset and split into training and test sets using random selection. The accuracy metric for the experiment is determined using the true detection rate.

For testing energy efficiency, a Raspberry Pi 3 B+ was chosen. Computational time of each step is tested. 10 trials are ran, and the mean execution time recorded for each step.

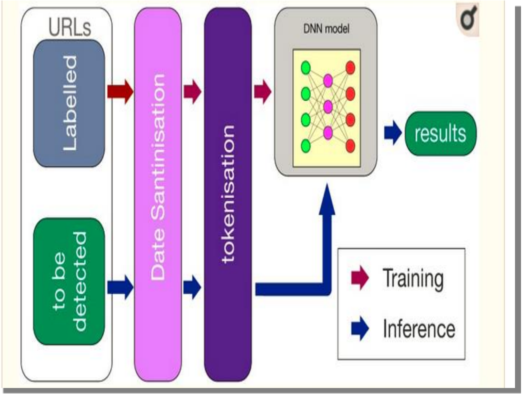
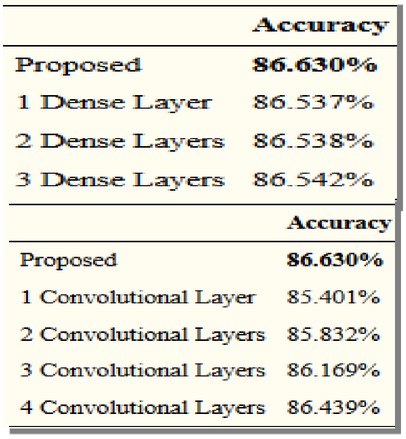


Figure 2 System structure of the proposed method, Bo Wei et. al.(2019)

After the experiments, the researchers concluded that their method has a 86.630% true detection rate. Furthermore, they noted that detection rate increases with more deep layers, convolutional layers and an embedding layer. For energy efficiency, evaluation of each URL request takes 110 ms, which they stated was sufficient for real-time detection. They then compared the proposed method with another method that utilizes word-level and character-level embedding, with the latter causing an out-of-memory error during testing.



	Accuracy
Proposed	**86.630%**
1 Dense Layer	86.537%
2 Dense Layers	86.538%
3 Dense Layers	86.542%

	Accuracy
Proposed	**86.630%**
1 Convolutional Layer	85.401%
2 Convolutional Layers	85.832%
3 Convolutional Layers	86.169%
4 Convolutional Layers	86.439%

Table 2 Accuracy of multiple dense and convolutional layers, Bo Wei et. al.(2019)

While the experiments are clearly detailed and documented, the source of the phishing URLs used was not mentioned, and legitimate URLs sourced from questionable sites, such as malwaredomains. The accuracy is fairly low, likely compromising for low energy requirements. The device used to test energy efficiency was appropriate, and clear comparisons with a similar model was made.

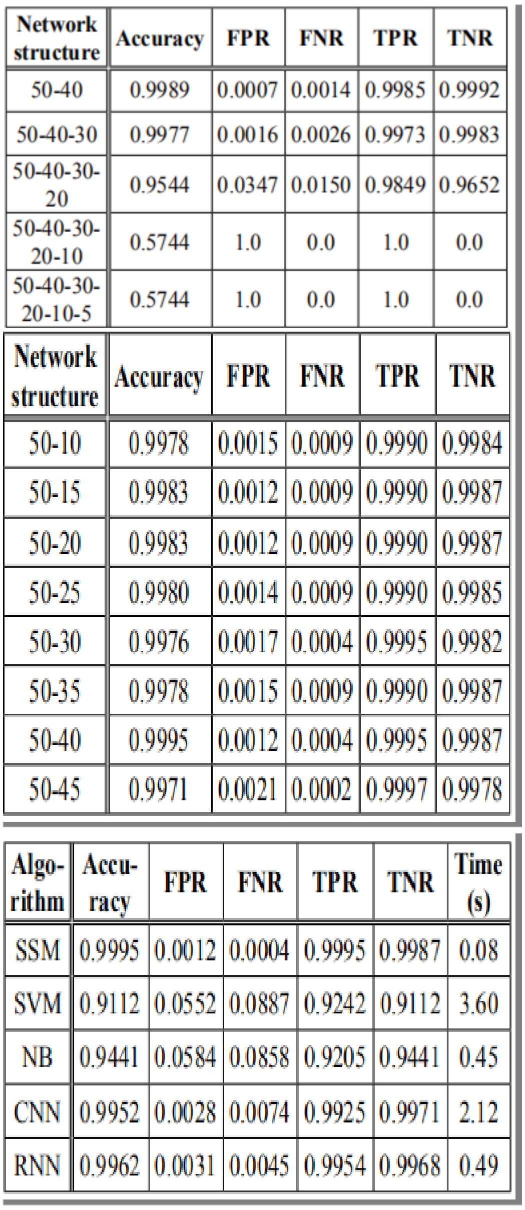
Jian Feng et. al.(2019) proposes a method that extracts features from website source code and third party services using deep learning methods to determine if a webpage is phish. They assert that the proposed method, SSM(SAE-Softmax model) is promising in phishing detection applications.

The researchers designed three sets of experiments to test the proposed model. The dataset used mixes 8848 legitimate URLs from Alexa and 11321 phishing URLs from PhishTank. The three experiments use the same hyperparameters except layer width. Each experiment has 4 iterations. The first experiment reduces layer width each iteration. The second experiment increases layer width each iteration. Both are then ran on new datasets, a classic dataset called GermanCredit and a dataset recreated from the original dataset. The third experiment changes nothing and compares the propounded method with existing detection methods.

In the first experiment, the researchers concluded that the best layer width was 50-40. In the second experiment, results are highest when layer width is 50-40. In the recreated dataset, the best width is 50-35. However, in the GermanCredit test, 3 layers with widths set to 20-12-4 gave the best results.

In experiment 3, the researchers found the proposed model shows better performance than existing detection methods. However, they note that this may not be correct for every situation, as not all possible hyperparameters were controlled.

All experiments were detailed well, allowing for clear reproducibility. The controls and manipulated factor were clearly stated. As the proposed model was tested using separate datasets and similar models, there is no bias in the results. However, GermanCredit's low detection rate suggests the method's effectiveness drops when too many factors are considered simultaneously.



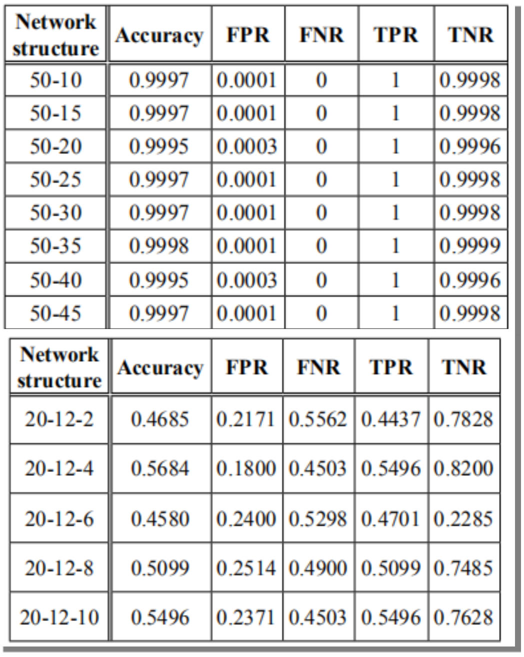
Network structure	Accuracy	FPR	FNR	TPR	TNR
50-40	0.9989	0.0007	0.0014	0.9985	0.9992
50-40-30	0.9977	0.0016	0.0026	0.9973	0.9983
50-40-30-20	0.9544	0.0347	0.0150	0.9849	0.9652
50-40-30-20-10	0.5744	1.0	0.0	1.0	0.0
50-40-30-20-10-5	0.5744	1.0	0.0	1.0	0.0

Network structure	Accuracy	FPR	FNR	TPR	TNR
50-10	0.9978	0.0015	0.0009	0.9990	0.9984
50-15	0.9983	0.0012	0.0009	0.9990	0.9987
50-20	0.9983	0.0012	0.0009	0.9990	0.9987
50-25	0.9980	0.0014	0.0009	0.9990	0.9985
50-30	0.9976	0.0017	0.0004	0.9995	0.9982
50-35	0.9978	0.0015	0.0009	0.9990	0.9987
50-40	0.9995	0.0012	0.0004	0.9995	0.9987
50-45	0.9971	0.0021	0.0002	0.9997	0.9978

Algo-rithm	Accu-racy	FPR	FNR	TPR	TNR	Time (s)
SSM	0.9995	0.0012	0.0004	0.9995	0.9987	0.08
SVM	0.9112	0.0552	0.0887	0.9242	0.9112	3.60
NB	0.9441	0.0584	0.0858	0.9205	0.9441	0.45
CNN	0.9952	0.0028	0.0074	0.9925	0.9971	2.12
RNN	0.9962	0.0031	0.0045	0.9954	0.9968	0.49

Table 3 First, second and third experiment results, Jian Feng et. al.,2019

The study's choice to extract features from not only URL but additional features is supported by Peng Yang et. al.(2019), who states that not only is URL occasionally uncertain for detection purposes, but phishers often imitate more than just URLs in phishing attacks .



Network structure	Accuracy	FPR	FNR	TPR	TNR
50-10	0.9997	0.0001	0	1	0.9998
50-15	0.9997	0.0001	0	1	0.9998
50-20	0.9995	0.0003	0	1	0.9996
50-25	0.9997	0.0001	0	1	0.9998
50-30	0.9997	0.0001	0	1	0.9998
50-35	0.9998	0.0001	0	1	0.9999
50-40	0.9995	0.0003	0	1	0.9996
50-45	0.9997	0.0001	0	1	0.9998

Network structure	Accuracy	FPR	FNR	TPR	TNR
20-12-2	0.4685	0.2171	0.5562	0.4437	0.7828
20-12-4	0.5684	0.1800	0.4503	0.5496	0.8200
20-12-6	0.4580	0.2400	0.5298	0.4701	0.2285
20-12-8	0.5099	0.2514	0.4900	0.5099	0.7485
20-12-10	0.5496	0.2371	0.4503	0.5496	0.7628

Table 4 Recreated and GermanCredit dataset results, Jian Feng et. al.,2019

Yong Fang et. al.(2019) proffers a phishing email detection model grounded on a recurrent convolutional neural network (RCNN) model that concurrently models emails on every level. They aim to quickly determine an email's legitimacy with high accuracy.

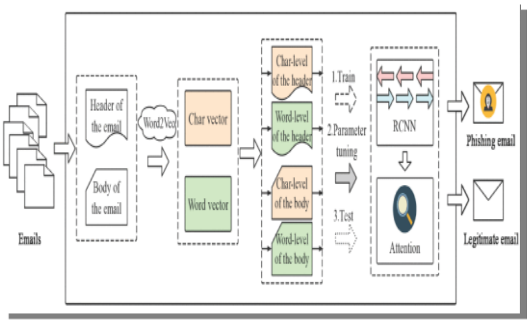
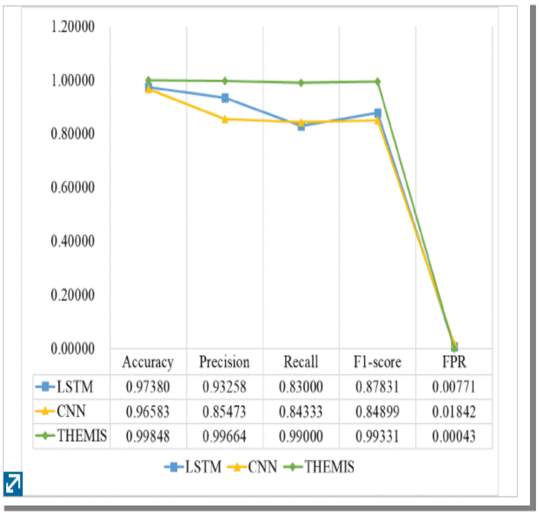


Figure 3 Framework for email classification, Yong Fang et. al.(2019)

They prepared a dataset by collecting phishing data from the First Security and Privacy Analytics Anti-Phishing Shared Task(IWSPA-AP). Legitimate emails are sourced from Wikileaks archives. The data is partially preprocessed, and only have emails sized 1MB or less. Only emails with headers are tested, and divided into a training-validation and testing set via stratified random sampling. 10-fold cross-validation is finally used on the sets.

Furthermore, noise is reduced by using the Python library “email” to eliminate empty spaces and divide the email into header and body. Both are segmented via word and character segmentation. Google Word2Vec is used to train the model’s character and vector models.

The training-validation set then trains the proposed model with 20 epochs. The model is verified after each epoch. A classification threshold of 0.79829 is set. Emails higher than the threshold are considered phish. Next, the testing set is inputted into the model for evaluation. A series of indicators are outputted and evaluated to determine performance. Lastly, CNN and LSTM based models are compared with the model using identical datasets.



	Accuracy	Precision	Recall	F1-score	FPR
LSTM	0.97380	0.93258	0.83000	0.87831	0.00771
CNN	0.96583	0.85473	0.84333	0.84899	0.01842
THEMIS	0.99848	0.99664	0.99000	0.99331	0.00043

Figure 4 Summary of experiment results, Yong Fang et. al.(2019)

The researchers concluded that the proposed model performed better than its counterparts. They note however, that their model is limited to emails with headers. It is also remarked that deep learning techniques are not fully interpretable, causing the exact reasons for differing performances to be uncertain.

The experiment’s controls were clearly stated, and the procedures were well documented. External factors, such as noise was reduced to the utmost. Furthermore, Yong Fang et. al.(2019) made clear comparisons with other existing models. The training regimen for the model was thorough with clear parameters. The results indicate the model can be expanded to direct messages on various applications.

Machine Learning is highly suitable for detecting phish with high accuracy, and the performance does not degrade much when other factors like energy consumption is taken into account. However, high training requirements may increase the time needed to implement this method.

2.3 Support Vector Machine

Support Vector Machine classification algorithm is a clustering algorithm that uses supervised learning and learning algorithms to analyze data classification and regression. M.A.Adebowale et. al.(2018) proposes a phishing detection scheme that uses integrated features in text, images and frames for web-based detection and protection via Adaptive Neuro-Fuzzy Inference System (ANFIS), an improved version of SVM. The Sugeno fuzzy model is used in conjunction with ANFIS.

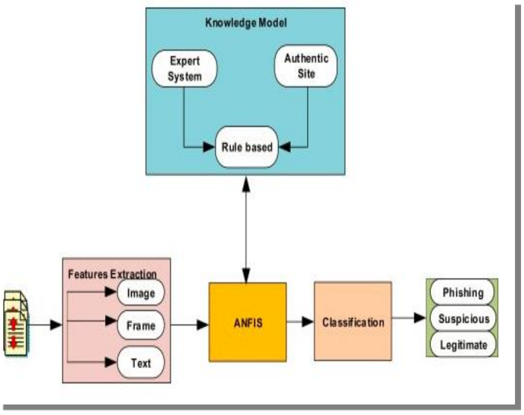


Figure 5 Proposed concept diagram, M.A.Adebowale et. al.(2018))

The dataset used for testing is taken from the University of California Irvine, the University of Huddersfield, PhishTank and the Anti-Phishing Working Group(APWG). These were then divided into 5 training sets and a testing set. The dataset contains “4,898 phishing websites, 1,945 suspicious sites and 6,157 legitimate websites”(M.A.Adebowale et. al.,2018).

The proposed scheme was implemented in MATLAB with parameters of 10 epochs, 0 error rate, hybrid optimisation method and 5-fold cross-validation. Supervised learning was used for training, and the 5 testing datasets divided into training and endorsement data. The datasets were then inputted arbitrarily.

Testing follows the same steps as training, with every inputted dataset used at least once. Performance was determined using the average error rate through dividing the error rate sum with data quantity. The performance is then compared with conventional K-nearest neighbour and Support Vector Machine classifications.

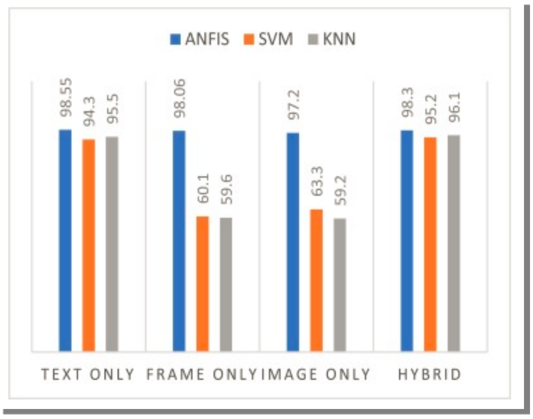


Figure 6 Experimental results, M.A.Adebowale et. al.(2018)

The researchers presented that based on Figure 7, the proposed scheme performs better than conventional SVM and KNN classifications, with a high accuracy rate of 98.3%. They assert that the proposed scheme can be more robust with higher element concentration.

The experimental procedures were thoroughly documented, allowing for reproducibility. The dataset was sourced from reputable sources, and a large sample size were used. The controls and the training procedures were clearly stated. Random inputting of testing and training sets ensured no bias was present throughout the experiment. The researchers also compared the performance of the proposed scheme with existing models. The propounded scheme could be used in threat prevention or ensuring websites follow required standards.

M.A.Adebowale et. al.(2018)'s assertions that basic SVM is insufficient, hence using ANFIS alongside Sugeno fuzzy model is supported by Routhu Srinivasa Rao et. al.(2019b), who gives evidence that conventional SVM alone may not adapt sufficiently to present-day phishing attacks.

3 Comparison of Methods

Following the critical analysis of every proposed method, Deep Learning based techniques offer the best results for detecting phishing attacks. Bo Wei et. al.(2019)'s deep learning algorithm achieved 86.63% accuracy with high energy efficiency. Jian Feng et. al.(2019)'s model featured 99.5% accuracy when tested, although it dropped to 56.84% when accounting for large quantity of factors. The propounded model by Yong Fang et. al.(2019) also shows high accuracy of 97.38%. All experiments utilized scientific rigor while being carried out and their hypothesis were scientifically sound. The large range in results suggests that deep learning is limited when considering other factors such as high factor quantity or energy efficiency.

Ozgur Koray Sahingoz et. al.(2018) showed that Random Forest combined with NLP features was highly effective, achieving an accuracy of 97.98%. However, as this was only tested against other classification algorithms in the same paper, the actual performance cannot be determined. Seow Wooi Liew et. al.(2019)'s usage of Random Forest, while effective could be improved using NLP features. It has however, been tested in a real-life scenario, meaning the results are applicable and consistent in real-life applications.

The proposed scheme using ANFIS by M.A. Adebowale et. al. (2018) has high accuracy when tested with no obvious drawbacks. While the accuracy is lower than Jian Feng et. al.(2019)'s model, the lack of drawbacks may make it a more attractive option.

Analysis shows that all 3 methods are highly effective at detecting phish, although further research is required to deal with the drawbacks and uncertainties in some methods which limit their application for real-life use.

4 Conclusions

In this research paper, various methods for accurately detecting phishing attacks have been critically analyzed. Of the evaluated methods, Deep Learning techniques that utilize feature extraction shows great promise due to high accuracy and low energy requirements, while being robust and performs in real-time. Classification models also show great promise, but as the evaluated models did not account for certain factors while testing, the veracity of their results are uncertain. Support Vector Machines are highly accurate as well, although Deep Learning techniques can surpass this method in most circumstances.

All experiments ran showed great consideration for scientific rigor, with clear controls and multiple experiments ran. While most experiments only considered phishing detection, Bo Wei et. al.(2019) showed that Deep Learning techniques can be energy saving at a cost to accuracy. Jian Feng et. al.(2019)'s experiments also showed that Deep Learning techniques' performance suffers if too many factors are considered at once. While Seow Wooi Liew et. al.(2019) did not give reasons for using Random Forest classification algorithm, they are backed up by Ozgur Koray Sahingoz et. al.(2018) who showed that it was the most effective classification algorithm. Furthermore, although Support Vector Machines alone are insufficient for reliably detecting phish, M.A. Adebowale et. al. (2018) and Routhu Srinivasa Rao et. al.(2019b) showed that by improving the SVM algorithm and/or combining it with other algorithms, SVMs are highly accurate at phishing detection with no immediate drawbacks.

To conclude, more research will be required to further improve the analyzed methods' accuracy without compromising on other factors. However, their high success rate in detecting phish allows for real-time use with good results.

References

Bo Wei, Rebeen Ali Hamad, Longzhi Yang, Xuan He, Hao Wang, Bin Gao, Wai Lok Woo, 2019, 'A Deep-Learning-Driven Light-Weight Phishing Detection Sensor', *Sensors,* Vol. 19, Iss. 19, pg. 4258

Carlo Marcelo Revoredo daSilva, Eduardo Luzeiro Feitosa, Vinicius CardosoGarcia, 2019, 'Heuristic-based strategy for Phishing prediction: A survey of URL-based approach', *Computers & Security,* Vol. 88, Art. 101613

Jian Feng, Lianyang Zou, Tianzhu Nan, 2019, 'A Phishing Webpage Detection Method Based on Stacked Autoencoder and Correlation Coefficients', *Journal of Computing and Information Technology,* Vol.27, Iss. 2, pg. 41-54

M.A. Adebowale, K.T. Lwin, E. Sánchez, M.A. Hossain, 2018, 'Intelligent web-phishing detection and protection scheme using integrated features of Images, frames and text', *Expert Systems With Applications,* Vol 115, pg. 300-313

Mahdieh Zabihimayvan, Derek Doran, 2019, 'Fuzzy Rough Set Feature Selection to Enhance Phishing Attack Detection', *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE),* No. 19046371

Ozgur Koray Sahingoz, Ebubekir Buber, Onder Demir, Banu Diri, 2018, 'Machine learning based phishing detection from URLs', *Expert Systems with Applications,* Vol. 117, pg. 345-357

Peng Yang, Guangzhen Zhao, Peng Zeng, 2019, 'Phishing Website Detection Based on Multidimensional Features Driven by Deep Learning', *IEEE Access,* Vol.7, pg.15196 - 15209

Routhu Srinivasa Rao, Alwyn Roshan Pais, 2019a, 'Jail-Phish: An improved search engine based phishing detection system', *Computers & Security,* Vol 83, pg. 246-267

Routhu Srinivasa Rao, Tatti Vaishnavi, Alwyn Roshan Pais, 2019b, 'PhishDump: A multi-model ensemble based technique for the detection of phishing sites in mobile devices', *Pervasive and Mobile Computing,* Vol.60, Art.101084

Seow Wooi Liew, Nor Fazlida Mohd Sani, Mohd. Taufik Abdullah, Razali Yaakob, Mohd Yunus Sharum, 2019, 'An effective security alert mechanism for real-time phishing tweet detection on Twitter', *Computers & Security,* Vol. 83, pg. 201-207

Tommy Chin, Kaiqi Xiong, Chengbin Hu, 2018, ‘Phishlimiter: A Phishing Detection and Mitigation Approach Using Software-Defined Networking’, *IEEE Access,* Vol 6, pg. 42516 - 42531

Yong Fang, Cheng Zhang, Cheng Huang, Liang Liu, Yue Yang, 2019, ‘Phishing Email Detection Using Improved RCNN Model With Multilevel Vectors and Attention Mechanism’, *IEEE Access,* Vol. 7, pg. 56329 - 56340