Computer Science & Information Technology 122

David C. Wyld
Dhinaharan Nagamalai (Eds)

Computer Science & Information Technology

8th International Conference of Security, Privacy and Trust Management (SPTM 2020)
June 13 ~ 14, 2020, Helsinki, Finland

AIRCC Publishing Corporation

Volume Editors

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

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

ISSN: 2231 - 5403
ISBN: 978-1-925953-20-0
DOI: 10.5121/csit.2020.100601- 10.5121/csit.2020.100611

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

Preface

The 8th International Conference of Security, Privacy and Trust Management (SPTM 2020) June 13 ~ 14, 2020, Helsinki, Finland, 6th International Conference on Image Processing and Pattern Recognition (IPPR 2020), 6th International Conference on Artificial Intelligence and Soft Computing (AIS 2020), 7th International Conference on Computer Science and Information Technology (CSIT 2020), 6th International Conference on Data Mining (DaMi 2020), International Conference on Advanced Machine Learning (AMLA 2020) was collocated with 8th International Conference of Security, Privacy and Trust Management (SPTM 2020). The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

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

The SPTM 2020, IPPR 2020, AIS 2020, CSIT 2020, DaMi 2020 and AMLA 2020 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically.

In closing, SPTM 2020, IPPR 2020, AIS 2020, CSIT 2020, DaMi 2020 and AMLA 2020 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the SPTM 2020, IPPR 2020, AIS 2020, CSIT 2020, DaMi 2020 and AMLA 2020.

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

David C. Wyld
Dhinaharan Nagamalai (Eds)

General Chair	**Organization**
David C. Wyld,	Southeastern Louisiana University, USA
Dhinaharan Nagamalai,	Wireilla Net Solutions, Australia

Program Committee Members

Ajantha Dahanayake,	Lappeenranta-Lahti University of Technology, Finland
Akhil Gupta,	Lovely Professional University, India
Amalina F. A. Fadzlah,	Universiti Pertahanan Nasional Malaysia
Amando P. Singun JR,	Higher College of Technology, Muscat, Oman
Amelia Badica,	University of Craiova, Romania
Anand Nayyar,	Duy Tan University, Vietnam
Azeddine WAHBI,	Hassan II University, Morocco
Bouhorma Mohammed,	Faculty of Science and Technology of Tangier, Morocco
Chandra Singh,	Sahyadri College of Engineering and Management, India
Chuanhe Huang,	Wuhan University, China
Cihat Cetinkaya,	Mugla Sitki Kocman University, Turkey
Cliff C. Zou,	University of Central Florida, USA
Desmond Bala,	Cranfield University, United Kingdom
Dirk Thorleuchter,	Fraunhofer INT, Germany
Elaachak Lotfi,	Abdelmalek Essadi University, Morocco
Estevan Gomez- Torres,	Universidad "UTE", Ecuador
Fatma Taher,	Zayed University, United Arab Emirates
Federico Tramarin,	University of Padova,Italy
Felix Yang Lou,	City University of Hong Kong,China
Franco Frattolillo,	University of Sannio, Italy
Geraldo Pereira Rocha Filho,	University of Brasilia, Brazil
Grammati Pantziou,	University of West Attica, Greece
Grigorios N.Beligiannis,	University of Patras, Greece
Hamid Ali Abed AL-Asadi,	Basra University, Iraq
Heba Elgazzar,	Morehead State University, USA
Imine Abdessamad,	University of Lorraine, France
Isaac Agudo,	University of Malaga, Spain
Jafar A. Alzubi,	Al-Balqa Applied University, Jordan
Jagadeesh HS,	APS College of Engineering (VTU), India
Jose Luis Verdegay,	University of Granada, Spain
Jui-Pin Yang,	Shih-Shien University, Taiwan
Karim MANSOUR,	University Salah Boubenider, Algeria
Klenilmar Lopes Dias,	Federal University of Minas Gerais, Brazil
M.K.Marichelvam,	Mepco Schlenk Engineering College, India
M.Suresh,	Kongu Engineering College, India
M.Vijayalakshmi,	Thiagarajar College of Engineering, India
Malleswara Talla,	Concordia University, Canada
Mohamed FAKIR,	Sultan Moulay Slimane University, Morocco

Mohamed Hamlich,	UH2C, ENSAM, Morocco
Mohamed Lehsaini,	Tlemcen University, Algeria
Morteza Alinia Ahandani,	University of Tabriz, Iran
Mourad Oussalah,	University of Nantes, France
Moussa Witti,	University of Geneva, Switzerland
Nadia Abd-Alsabour,	Cairo University, Egypt
Narendra V G,	Manipal Institute of Technology, India
Neveen I. Ghali,	Future University, Egypt
Nor Shahida Mohd Jamail,	Prince Sultan University, Saudi Arabia
Onur Gunlu,	Technical University of Berlin, Germany
Osama Rababah,	University of Jordan, Jordan
Othmane Alaoui Fdili,	Cady Ayyad University, Morocco
Picky Butani,	Shubh Solutions LLC, USA
Pierre Sylvain Iloga Biyik,	University of Maroua, Cameroon
Prabhat Mahanti,	University of New Brunswick,Canada
Pramod Kumar,	Manipal Institute of Technology, India
Qasim Zeeshan Ahmed,	University of Huddersfield, United Kingdom
Rachid LATIF,	Ibn Zohr University, Morocco
Ramgopal Kashyap,	Amity University Chhattisgarh, India
Rosalba Cuapa Canto,	Universidad Autonoma de Puebla, Mexico
Shahid Ali,	AGI Education Ltd, New Zealand
Shirish Patil,	Lead Enterprise Data Architect, Sitek Inc, USA
Simanta Shekhar Sarmah,	Alpha Clinical Systems Inc, USA
Smitha N. Pai,	Manipal Institute of Technology, India
Solale Tabarestani,	Florida International University, USA
Solomiia Fedushko,	Lviv Polytechnic National University, Ukraine
Sos Agaian,	College of Staten Island and the Graduate Center, USA
Subramanian Selvakumar,	NIT Tiruchirappalli, India
Thayalini Majureshan,	University of Westminster, Sri Lanka
Tonghan Wang,	East China University of T technology, China
Uduak Augustine Umoh,	University of Uyo, Nigeria
Uduak Umoh,	University of Uyo, Nigeria
Wael Al Zoubi,	Balqa Applied University, Jordan
wael jumah alzyadat,	Isra university, Jordan
Wei Hong Lim,	UCSI University, Malaysia
Weili Zhang,	research scientist eBay Inc., USA
Wesam M. Jasim,	University of Anbar, Iraq
Xuechao Li,	Auburn University USA
Yung Gi Wu,	Chang Jung Christian University, Taiwan

Technically Sponsored by

Computer Science & Information Technology Community (CSITC)

Artificial Intelligence Community (AIC)

Soft Computing Community (SCC)

Digital Signal & Image Processing Community (DSIPC)

Organized By

Academy & Industry Research Collaboration Center (AIRCC)

TABLE OF CONTENTS

8th International Conference of Security, Privacy and Trust Management (SPTM 2020)

6th International Conference on Image Processing and Pattern Recognition (IPPR 2020)

6th International Conference on Artificial Intelligence and Soft Computing (AIS 2020)

7th International Conference on Computer Science and Information Technology (CSIT 2020)

6th International Conference on Data Mining (DaMi 2020)

International Conference on Advanced Machine Learning (AMLA 2020)

FCNNMD: A NOVEL FUSION METHOD BASED ON CONVOLUTIONAL NEURAL NETWORK FOR MALWARE DETECTION

Jing Zhang[1] and Yu Wen[2]

[1]Institute of Information Engineering, CAS & School of Cyber Security,
University of Chinese Academy of Sciences, Beijing, China
[2]Institute of Information Engineering, Chinese Academy of Sciences,
Beijing, China

ABSTRACT

Malicious software are rampant and do great harm. The present mainstream malware detection technology has many disadvantages, such as high labour cost, large system overhead, and inability to detect new malware. We propose a novel fusion method based on convolutional neural network for malware detection (FCNNMD). For the sample imbalance problem faced by the convolutional neural network malware detection method, the non-malicious sample is added by means of generating anti-network generation, etc., to achieve the same number as the malicious sample. For the problem of low accuracy of single model detection, high false positive rate and false negative rate, a malware detection model is constructed by means of model fusion. The model combines four classical convolutional neural network structures. Experiments show that this method can effectively improve the accuracy and robustness of the model. Our method does not need actual running software and has high detection efficiency.

KEYWORDS

Malware Detection, Grayscale Image, Convolutional Neural Networks, Model integration

1. INTRODUCTION

In recent years, the Internet security reports of major security manufacturers at home and abroad show that viruses, worms, trojans and other malicious software are ram-pant and do great harm. In addition, the number of malware and variant malware is still growing rapidly, and the network security situation is still very serious. The present mainstream malware detection technology has many disadvantages, such as high labour cost, large system overhead, easy to be bypassed by malware, and inability to detect new malware.

We propose a novel fusion method based on convolutional neural network for malware detection (FCNNMD). The basic idea is to use the grayscale image generation algorithm to convert the executable file into grayscale images firstly, and then use the convolutional neural network to build a grayscale image classification model. By classifying grayscale images, the purpose of detecting malware is achieved indirectly.

The detection method of malware based on convolutional neural network cannot only reduce human cost and system overhead, improve detection efficiency, but also detect new types of

David C. Wyld et al. (Eds): SPTM, IPPR, AIS, CSIT, DaMi, AMLA - 2020
pp. 01-16, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100601

malware. In addition, the detection method of malware based on convolutional neural network does not need to rely on artificial features. The main work of this thesis includes:

● For the sample imbalance problem faced by the convolutional neural network malware detection method, the non-malicious sample is added by means of generating anti-network generation, etc., to achieve the same number as the malicious sample. Experiments show that this method can effectively improve the accuracy of the model, reduce the false positive rate and false negative rate of the model.

● For the problem of low accuracy of single model detection, high false positive rate and false negative rate, a malware detection model is constructed by means of model fusion. The model combines four classical convolutional neural network structures. Compared with the single model detection, the accuracy of the model is further improved, and the false alarm rate and the false negative rate are further reduced.

● Aiming at the shortcomings of high complexity and low detection efficiency of model fusion method, a detection model with low complexity and high detection efficiency is designed and implemented. The accuracy, false positive rate and false negative rate of the model are passed. The model obtained by the model fusion method is basically the same, but the detection efficiency is improved by 3.15 times.

The malware detection method proposed in this paper can achieve an accuracy rate of about 97%, which is about 18% higher than the detection method based on texture fingerprint for malicious code variants. The detection method of malware based on convolutional neural network does not need actual running software, has high detection efficiency, does not rely on the construction of feature database, etc., and can greatly reduce labour cost and maintenance cost.

The rest of this paper is organized as follows: we briefly review the related work in Section 2, and we describe the proposed method in Section 3. The experimental results and discussion are presented in Section 4. Finally, we give conclusion and future work in Section 5.

2. RELATED WORK

In essence, malware detection technology is divided into two categories: static analysis and dynamic analysis. At present, most security manufacturers still rely on signature for malware detection, which belongs to static analysis [1]. This method can efficiently detect known malware, but it cannot detect new malware [2-4]. Dynamic analysis [5] is mainly based on the behaviour of software to detect malware. It usually need to actually run software to determine whether a software is malware, the most common practice is to rely on sandboxing programs. This detection method [6] has the disadvantages of high overhead, easy to be discovered by the defence detection mechanism of malware, and easy to be bypassed by malware.

Detection techniques based on machine learning and deep learning can detect new malware [34-36]. The detection method based on machine learning [7-9] mainly starts with the text structure of malware and extracts artificial features from many angles. [10] mainly identified malware by clustering, it first analysed the API of a large number of malware calls and used these API to form the dataset, then it extracted the subject from this dataset to cluster, and identified malware similar to known malware in the dataset. This method is obviously powerless against new malware [16]. Hyrum S et al. constructed an open source, large-scale PE format file data set, modeled and analysed the data set with the methods of machine learning and deep learning. The accuracy is about 95% [11]. The accuracy of detection method based on machine learning is usually low, and the number of samples used is very small, mostly between hundreds and thousands, and the robustness of the model is not high.

Deep learning [18] outperforms machine learning in many fields. In recent years, many researchers at home and abroad have shifted their research focus to malware detection based on deep learning [12-14]. Most of these methods are based on convolutional neural network [19-21] and recurrent neural network. Among them, the detection methods based on recurrent neural network structure are mostly based on software API call sequences to construct data sets. The accuracy of this kind of detection method is about 97%, but this kind of detection method has the problem of low detection efficiency, at the same time, because the sample is difficult to collect, most of them have the problem of small number of samples and low robustness of the model. The approach of [15] is to obtain the API call sequence of the software and consider the return value of each API. Finally, the classification model is established using the recurrent neural network. In [16], it obtains the API call sequence of the software and the API call sequence of the C language library, and then uses the recurrent neural network and echo state network to establish the classification model. The approach of [33] converts binary executable file into grayscale image, then extracts texture fingerprint based on grayscale image, it proposes a malicious code variant detection method based on texture fingerprint. Edward Raff et al. use convolutional neural network model that based on the bytecode sequence of the whole executable file, and a relatively high detection rate is also obtained, but the main problem is that the hardware requirements are too high, the computation is very large, and the model is relatively simple [17].

We propose a novel fusion method based on convolutional neural network for malware detection (FCNNMD). This method does not need to construct the malicious code feature library, but trains a grayscale image classification model through convolution neural network, which greatly reduces the labour cost and maintenance cost. Importantly, our model can detect new viruses. Compared with other methods of using machine learning to detect malware, the method does not need to rely on artificial features, but rely on convolutional neural network to automatically slave image species to extract features. This method does not need the actual running software, the detection efficiency is high, and the accuracy rate is about 97%. At the same time, samples are relatively easy to collect. This method can improve the accuracy and robustness of the model based on large sample training model.

3. PROPOSED METHODOLOGY

The overall process of our method consists of two parts: the classification part and the detection part (see Figure 1). First, we need to collect a large number of executable files containing malicious programs and a large number of executable files that do not contain malicious programs to build an executable file set. Then all files in the executable file set are converted to grayscale images using grayscale image generation algorithm to get a grayscale image set, and then we use convolutional neural network to train a classification model based on this grayscale image set, which can then be used in the detection process to detect whether an executable file contains malicious programs.

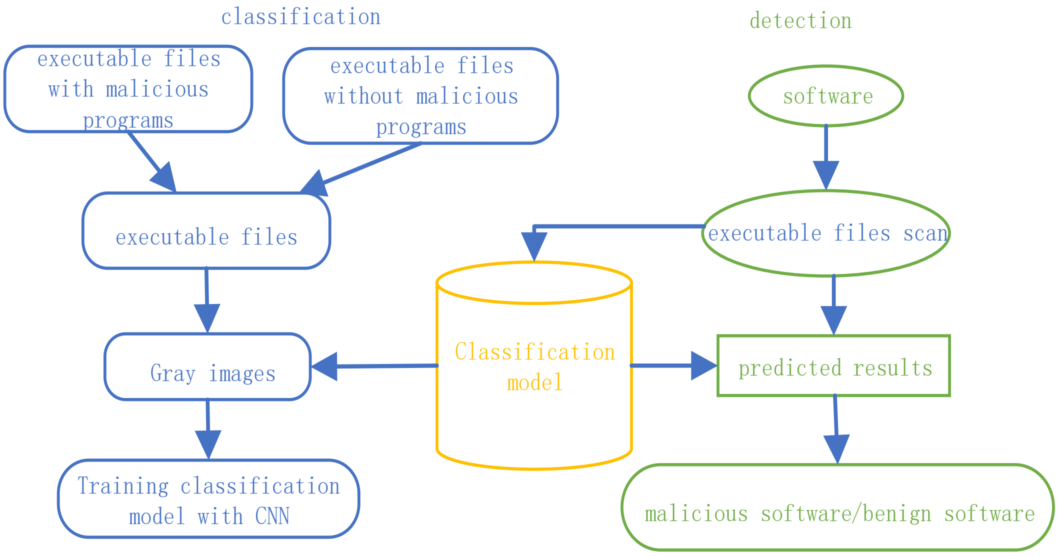

classification
executable files with malicious programs
executable files without malicious programs
executable files
Gray images
Training classification model with CNN
Classification model
detection
software
executable files scan
predicted results
malicious software/benign software

Figure 1. The overall flow of malware detection based on convolutional neural networks

3.1 Grayscale Image Conversion

Among the malware detection methods based on convolutional neural network, the executable file should be converted to grayscale image first, then the malware is detected based on the texture features in the grayscale image. B2M grayscale image generation algorithm is usually used when converting an executable file to a grayscale image [33]. The algorithm is to read any binary file every 8 bits as an unsigned integer (range between 0~255), and then set a fixed row width directly according to experience, so that after the whole file reading is finished, a two-dimensional array of unsigned integers will be obtained. Grayscale images of executable files containing the same family of malicious programs have similar texture features. Grayscale images of different executable files contained in the same software have similar texture features. The previous work does not specify the height and width of the grayscale image. This paper proposes that the form of the grayscale image should meet the following two conditions:

● The width of grayscale images should not be greater than the input width of convolutional neural networks. This ensures continuity of the header information and information of the executable file as much as possible while training the model. The header of the executable file is a key part of judging whether an executable file contains malicious programs. Experiments show that the accuracy can be increased about 2 percentage points by ensuring that the header information of the executable file is not lost.

● Grayscale images corresponding to executable files of similar size should also have similar image sizes. Executable files containing the same family of malicious programs are usually similar in size, because these executable files have similar functions, there is a great similarity between files. When converting these executables into grayscale images, it is ensured that the grayscale images have similar sizes, which ensures that the corresponding grayscale images of these executables have similar textures.

3.2 Sample Imbalance

In order to solve the problem of sample imbalance, after we get the grayscale image set, we adopt the following two ways to increase the number of grayscale image corresponding to the executable file without malicious program in the grayscale image set:

By cutting, flipping, flipping, etc. It is a more common method of data enhancement and can play a role on most data sets;

We use the grayscale image corresponding to the executable file without malicious program to train a generative adversarial network [32] model that can generate the grayscale image corresponding to the executable file without malicious program. First, we create a set of executable files that do not contain malicious programs, and convert all executable files into grayscale images to get a grayscale image set, then we select the image from the grayscale image set and input it into discriminator to let the discriminator judge whether the picture is generated by the generator. In addition, we randomly generate a set of noise data and input noise data into the generator to generate the image, then we input generated image into the discriminator to let the discriminator judge whether the image is generated by the generator. Finally the model adjusts the parameters in the generator and discriminator according to the judgment result of the discriminator, and continuously improve the generator's generating ability and discriminator's discriminant ability. After training, we use the generator to generate images to add to the dataset. Figure 2 shows the images of generator generated.

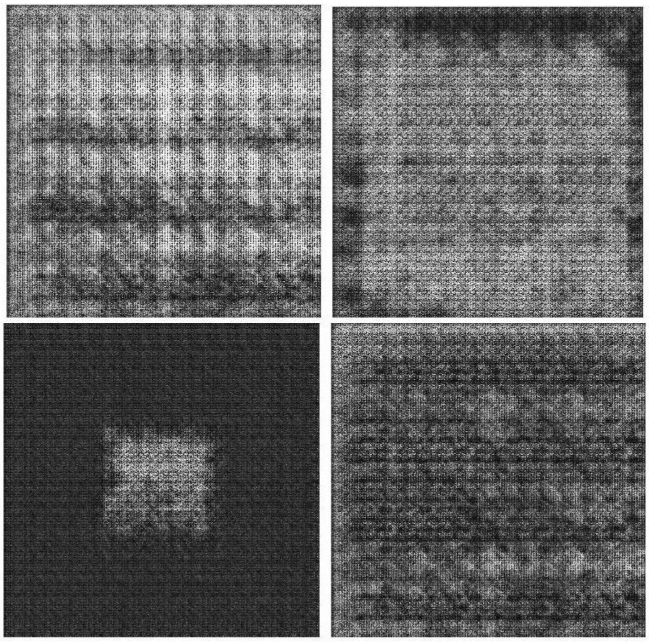
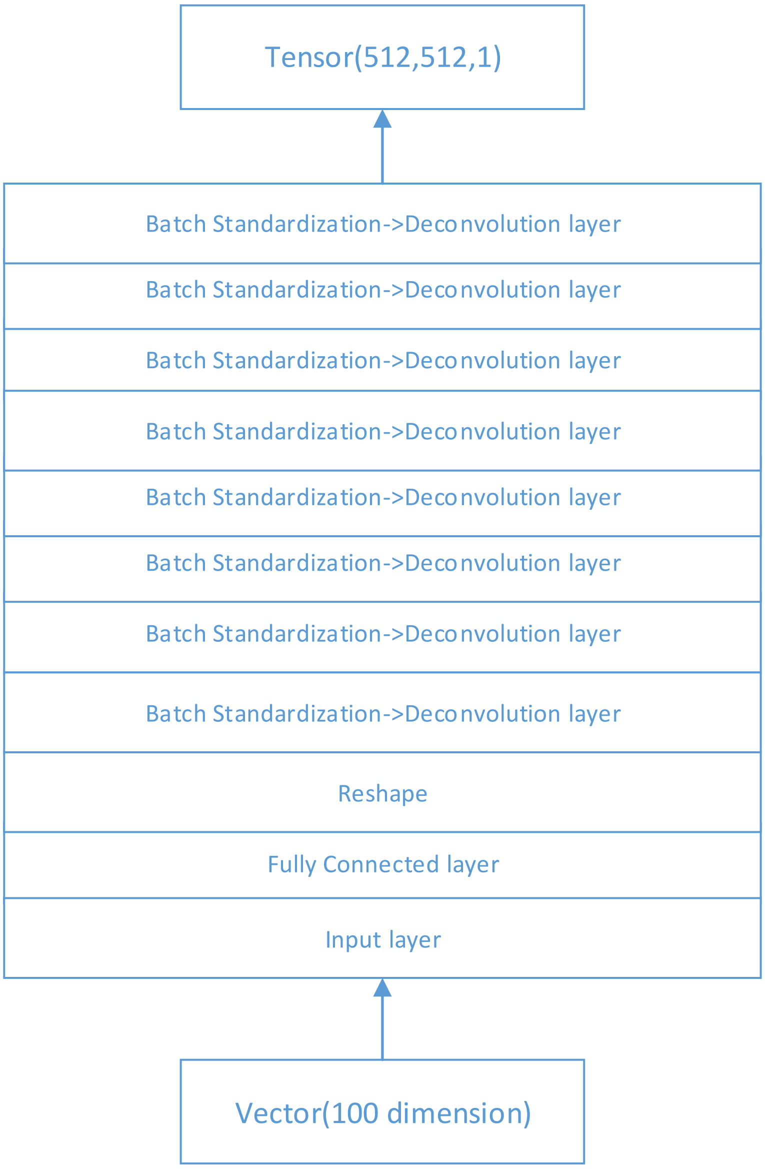


Figure 2. The images of generator generated

The structure diagram of the generator and discriminator in the generative adversarial network is shown in Figure 3.


Tensor(512,512,1)
Batch Standardization->Deconvolution layer
Batch Standardization->Deconvolution layer
Batch Standardization->Deconvolution layer
Batch Standardization->Deconvolution layer
Batch Standardization->Deconvolution layer
Batch Standardization->Deconvolution layer
Batch Standardization->Deconvolution layer
Batch Standardization->Deconvolution layer
Reshape
Fully Connected layer
Input layer
Vector(100 dimension)

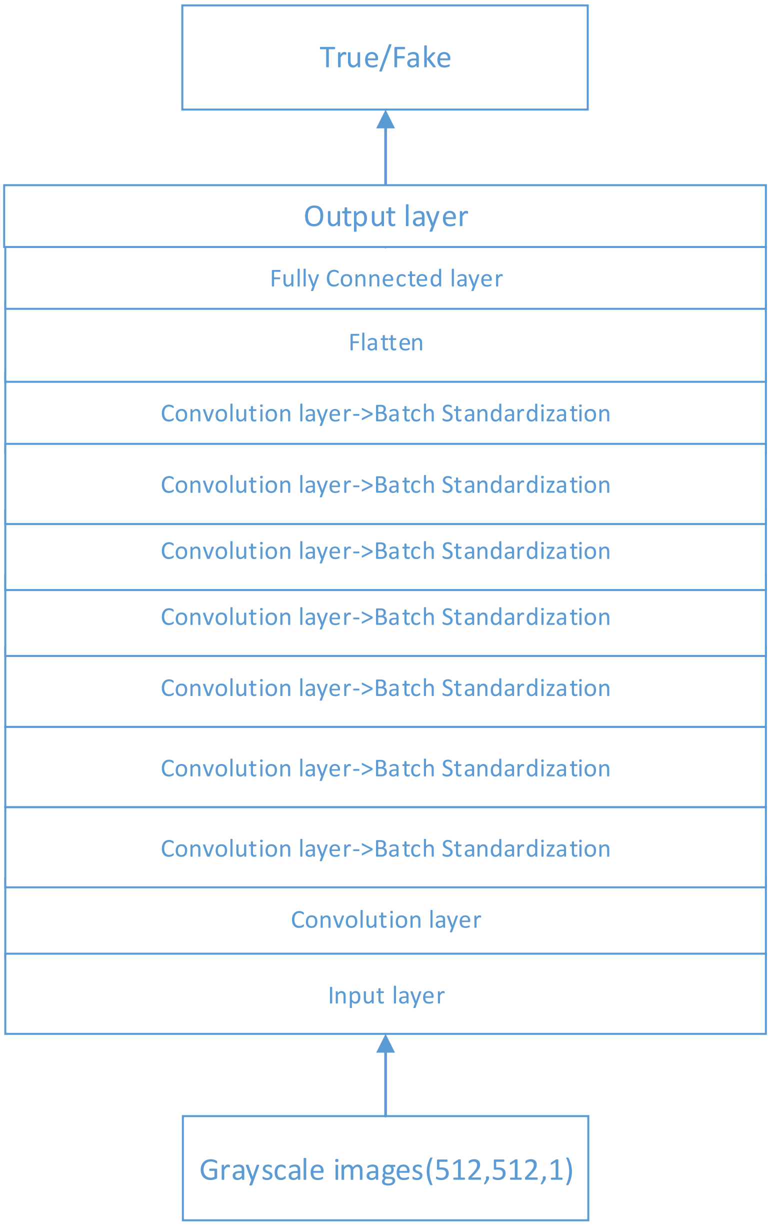

True/Fake
Output layer
Fully Connected layer
Flatten
Convolution layer->Batch Standardization
Convolution layer->Batch Standardization
Convolution layer->Batch Standardization
Convolution layer->Batch Standardization
Convolution layer->Batch Standardization
Convolution layer->Batch Standardization
Convolution layer->Batch Standardization
Convolution layer
Input layer
Grayscale images(512,512,1)

Figure 3. Generator structure diagram (a) Discriminator structure diagram (b)

Through the above two ways, we increase the original grayscale image set without malicious program executable corresponding to the number of grayscale image set to get the new grayscale image set, we delete part that contains malicious program executable file corresponding to the grayscale image, so that the number of two kinds of images in the data set is the same.

3.3 Model fusion

The model based on VGG16 [25], Inception-v3 [26-29], Xception [31] and ResNet50 [22] [30] is analyzed in detail by experiments (section 4). The following conclusions were drawn: VGG16 can effectively improve the accuracy of the model and reduce the false alarm rate; deep separable convolution layer in Xception can effectively reduce the failure rate of the model; the model with the least time overhead is the Incepiton-v3 model and its detection efficiency is highest. Based on the above conclusions, this paper designs and implements a complete model from bottom up, as shown in Figure 4.

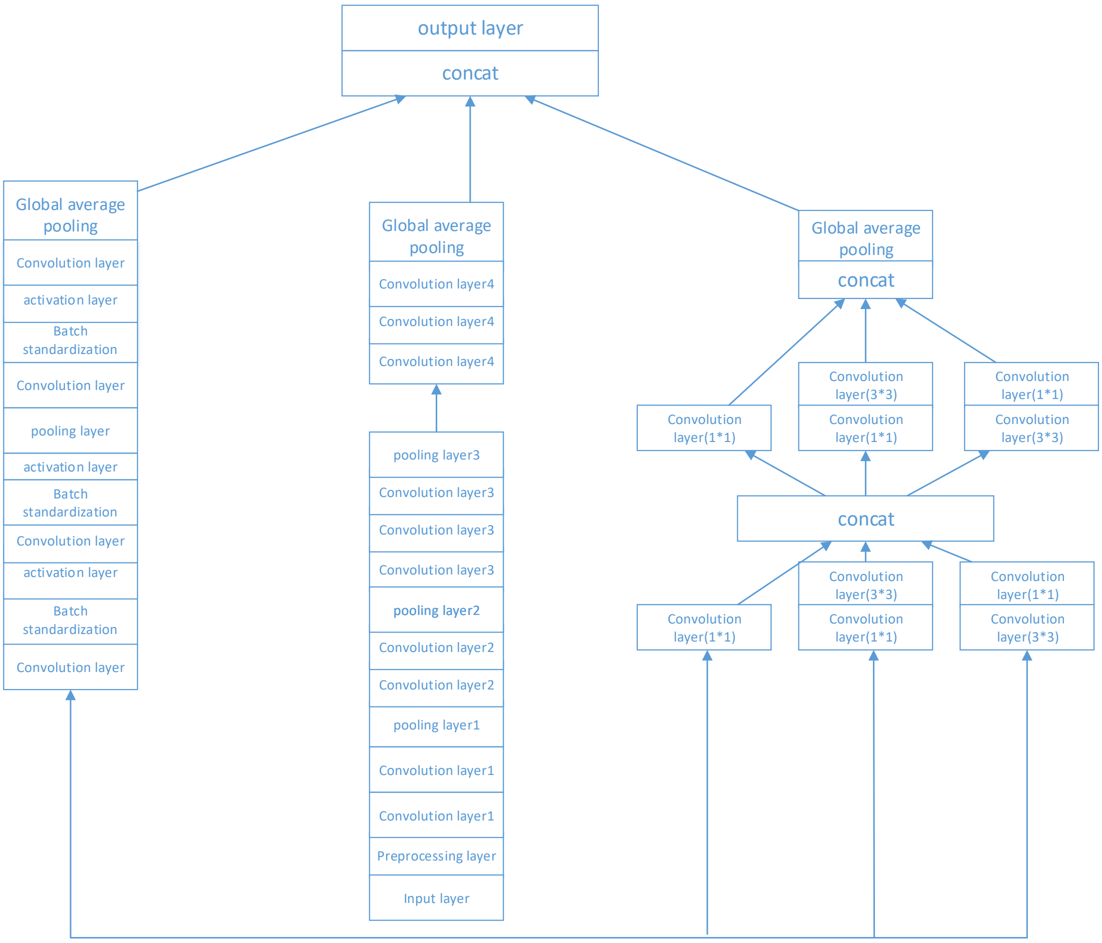


Figure 4. Custom model structure diagram

The input of the model is a grayscale picture, and a preprocessing layer is added after the input layer. The preprocessing layer is responsible for filling the grayscale picture with height or width less than 512 pixels, while scaling all the pixel values in the picture to between 0 and 1.

The main part of the model is mainly composed of 4 convolution pooling modules, each convolution pooling module is composed of multiple convolution layers and 1 pooling layer, the size of the convolution kernel of all convolution layers is 3*3, the step size is 1 and the padding mode is same padding. This part mainly takes into account that the accuracy and false positives of VGG16 in all single models are the lowest, so the trunk part adopts a structure similar to the VGG16 to ensure that the model has high accuracy and low false alarm rate.

The left branch of the model is mainly composed of 4 deep separable convolution layers, it mainly takes into account that the missing rate of Xception in all single models is the lowest, and then it is speculated that the Xception deep separable convolution structure has a role in reducing the missing rate of the model. At the same time, to reduce the complexity of the model, multiple deep separable convolution layers are used as branches to further extract features based on the first three convolution pooling modules of the trunk part, it is expected to reduce the missing rate of the model. Finally, the batch standardization layer is added to the left branch to accelerate the model convergence and prevent overfitting.

The right branch of the model contains multiple branch structures, it mainly draw lessons from the idea of Inception Module. First, 1*1 convolution is used in multiple branches .1*1 convolution can organize information across channels. Secondly, after each 1*1 convolution operation, the convolution results are nonlinearly calculated using the ReLU activation function. More nonlinear operations are introduced for the model to enhance the fitting ability of the model. Using this structure can also further increase the width of the model and the adaptability to the scale of the network, so there is multi-scale information inside. Moreover, the efficiency of this structure is relatively high. The whole right side branch of the model is mainly responsible for extracting more features in a more efficient way to further improve the performance of the model.

Finally, both the trunk part and the left and right side branches obtain a one-dimensional vector through a global average pooling layer. These three one-dimensional vectors are connected, then the grayscale image through the output layer is the probability of the grayscale image corresponding to the executable file containing the malicious program.

4. EXPERIMENTS

4.1. Experimental Settings

4.1.1. Datasets

We collect a large number of executable files containing malicious programs and executable files without malicious programs through various channels. Executable files that do not contain malicious programs are mainly from the mainstream Windows operating system, and executable files that contain malicious programs are mainly from sites that specifically collect malware. The composition of the data sets is shown in Figure 5:

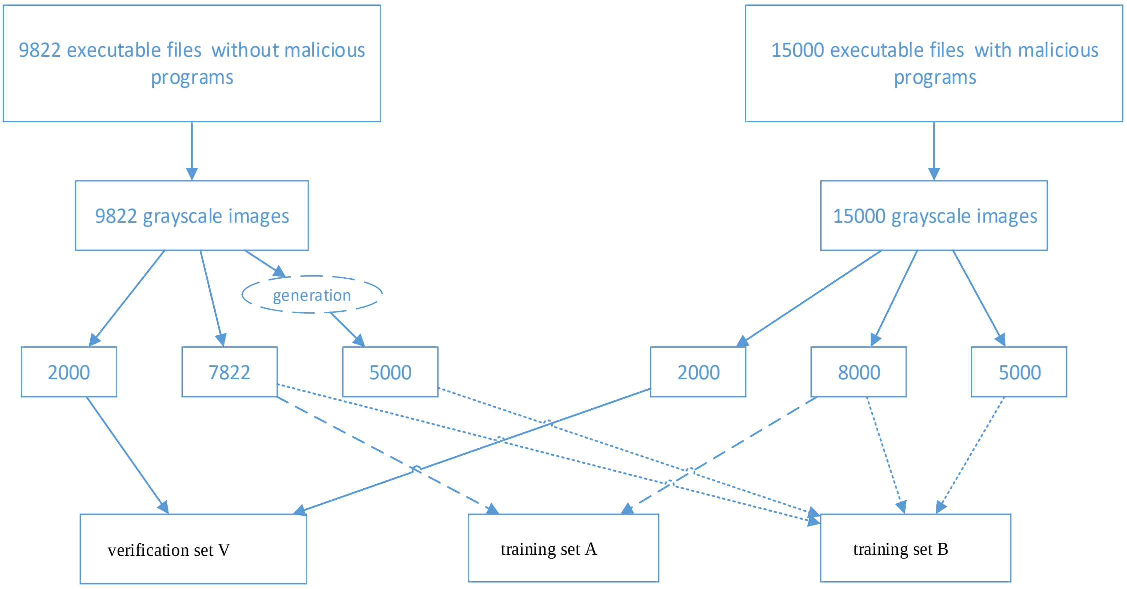


Figure 5. The division of data sets

Figure 5 shows that a total of 9822 executable files without malicious programs and 15000 executable files containing malicious programs participated in the construction of the data set. First, all the executable files are transformed into grayscale images using the improved B2M grayscale image generation algorithm. Then, based on 9822 grayscale images corresponding to executable files that do not contain malicious programs, 5000 grayscale images are generated using the method mentioned above. Then 9822 grayscale images of executable files that do not contain malicious programs were divided into two parts, one containing 2000 and the other containing 7822. Meanwhile, 15000 grayscale images of executable files containing malicious programs are divided into 3 parts. Finally, the verification set V, training set A and training set B are constructed according to the graph.

4.1.2. Metrics

For evaluating the performance of each model on the verification set V, the main evaluation indexes are accuracy, false alarm rate, missing report rate and time overhead. Time overhead is the time taken by each model to evaluate 4000 grayscale images V the validation set.
Accuracy:

$$Acc\text{=}\frac{TP+TN}{TP+TN+FN+FP}\times 100\% \tag{1}$$

False alarm rate:

$$False\ alarm\ rate=\frac{FN}{TP+FN}\times 100\% \tag{2}$$

Missing report rate:

$$Missing\ report\ rate=\frac{FP}{FP+TN}\times 100\% \tag{3}$$

4.1.3. Detailed Implementation

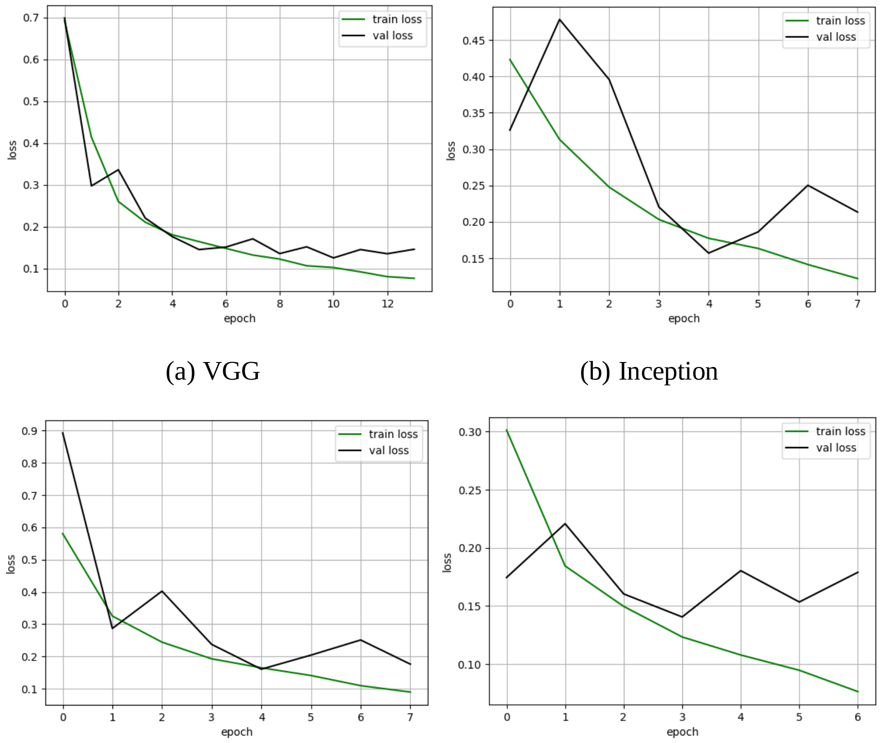
In this paper, the operating system used in the experiment is Linux system (Ubuntu 16.04), the width and height of grayscale images are defined as 512 pixels. DCGAN [32] is adopted in this paper. In our model, batch size is 8, epoch number is 20, and Adam is the optimization algorithm used. In order to prevent overfitting, the early stop strategy is also used.

4.2. Performance Comparison

4.2.1. Single Model Comparison

VGG, Inception, ResNet, Xception are four classical convolutional neural network structures. Most of the convolutional neural network structures are based on the further combination and improvement of these network structures, so in this paper, we first select these four classical convolutional neural networks to evaluate the performance, then we try to construct a better model based on the performance of these four convolutional neural networks.

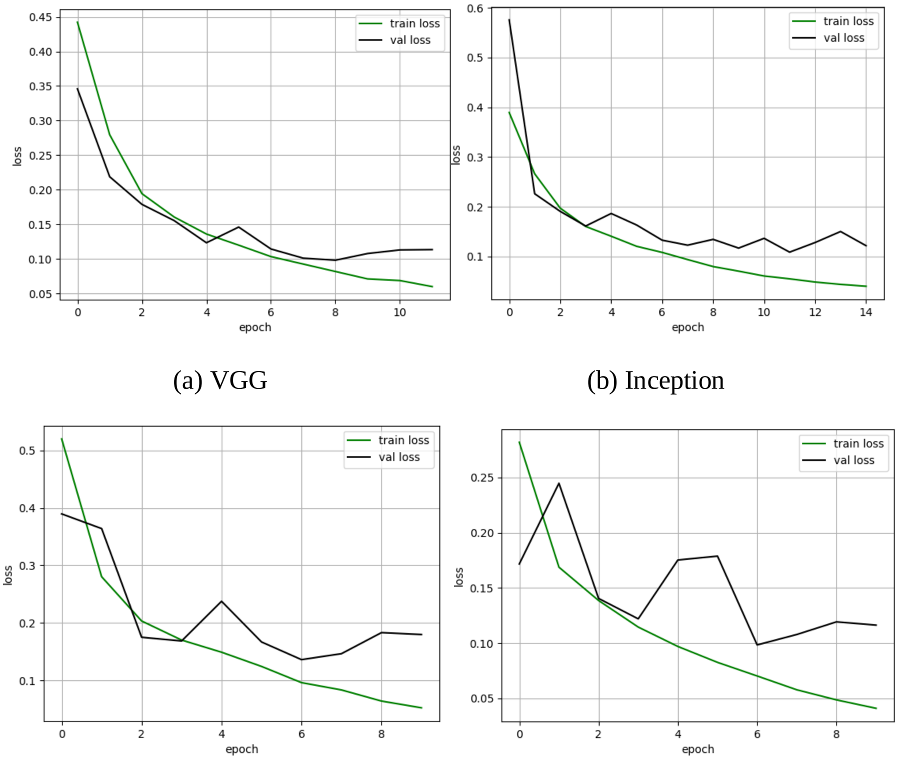
We have modified the network structure by adding a preprocessing layer before the input layer, which will fill images with height and width less than 512 pixels in the form of nearest neighbor interpolation. The preprocessing layer will also scale all pixel values of images to between 0 and 1, which can accelerate model convergence.



(a) VGG (b) Inception

(c) ResNet (d) Xception

Figure 6. Loss curve (training set A, verification set V)



(a) VGG (b) Inception

(c) ResNet (d) Xception

Figure 7. Loss curve (training set B, verification set V)

Table 1. The evaluation result of each model on training set A and verification set V

Model	**accuracy**	**false alarm rate**	**missing report rate**	**time overhead**
VGG16	95.92%	3.46%	4.59%	58s
ResNet50	94.40%	6.26%	4.94%	54s
Xception	95.16%	6.87%	2.80%	62s
Inception-v3	94.20%	5.95%	5.65%	47s
average	94.92%	5.64%	4.50%	/

Table 2. The evaluation result of each model on training set B and verification set V

Model	**accuracy**	**false alarm rate**	**missing report rate**	**time overhead**
VGG16	96.85%	2.80%	3.47%	58s
ResNet50	95.34%	5.02%	4.30%	54s
Xception	96.64%	3.97%	2.75%	62s
Inception-v3	96.30%	3.80%	3.42%	47s
average	96.28%	3.90%	3.49%	/

The following conclusions can be drawn by observing the above figure and tables:

● The evaluation results of the four models on the validation set V using the training set B are better than that using the training set A. This proves that the scheme used in this paper can improve the model performance and enhance the generalization ability of the model. In order to solve the problem of sample imbalance, we use the generative adversarial network generation network as one of the ways to expand the data set, and the generative adversarial network generates samples with similar distribution but not exactly the same as the original sample, so we can consider using the generative adversarial network to further expand the data set.

● For VGG16, the false alarm rate is the lowest and the accuracy is the highest, which shows that the VGG16 structure can effectively improve the accuracy of the model and reduce the false alarm rate of the model;

● Whether using training set A or training set B training model, the underreporting rate of Xception validation set is obviously lower than that of other models, which indicates that the deep separable convolution layer in Xception can effectively reduce the missing rate of the model.

● The model with the least time overhead is the Incepiton-v3 model, which shows that the Inception-v3 model is the most efficient.

4.2.2. Fusion Model Comparison

First, we design a simple model fusion method. Using previously trained network structures VGG16, Xception, ResNet50 and Inception-v3 to extract features, a one-dimensional vector of a certain length can be obtained through a global average pooling operation after passing through the intermediate structure of each network, then all one-dimensional vectors are connected into a long vector as the input of the fully connected layer. The output layer outputs the probability that if the grayscale image corresponds to the executable file containing the malicious program.

The model FCNNMD was finally trained on the training set B and the model performance was evaluated on the validation set V.

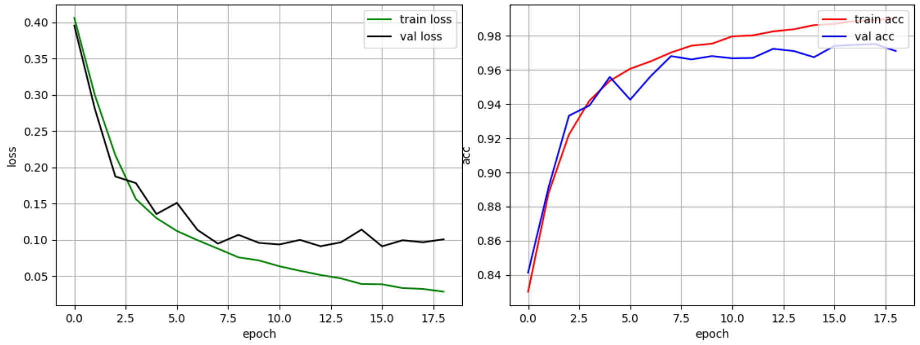


Figure 8. Loss and accuracy curve of custom model (training set B, verification set V)

The left figure of Figure 8 shows that with the increase of the number of epoch, the loss value of the model on the training set B and the verification set V shows a downward trend. From the right figure, we can see that the accuracy of the model on the training set B and the verification set is increasing with the increase of the number of epoch. In the 16th epoch, the model has the lowest loss value and the highest accuracy.

Table 3. The evaluation result of fusion model on verification set V

Metric	training set A	training set B
accuracy	96.65%	97.70%
false alarm rate	3.21%	3.06%
missing report rate	3.49%	1.54%
time overhead	230s	230s

The comparison with the results of single model shows that the performance of the model can be further improved by model fusion (Table 3), and the generalization ability of the model can be enhanced. The model obtained by model fusion has higher accuracy rate, lower false alarm rate and missing alarm rate.

Table 4. The evaluation result on training set B

Metric	Simple Fusion	FCNNMD
accuracy	97.70%	97.8%
false alarm rate	3.06%	3.05%
missing report rate	1.54%	1.35%
time overhead	230s	73s

As can be seen from Table 4, the simple fusion model has high complexity and low detection efficiency, and the time overhead of the model is basically equal to the sum of the time overhead of the four single models. FCNNMD complexity is greatly reduced, and the detection efficiency of the model is greatly improved, and the overall improvement is 3.15 times. The reason why the efficiency of model detection is greatly improved is that the number of parameters in the model is greatly reduced and the model is simplified by optimization.

In our method, when adjusting parameters in reverse, only the parameters in full connection layer are adjusted without changing the parameters in each network structure, which weakens the effect of model fusion. Sufficient hardware resources allow all models to be loaded into memory. When adjusting parameters in reverse, the parameters of some layers in each network structure can be fine-tuned according to the results of model fusion.

5. CONCLUSION AND FUTURE WORK

In this paper, we propose a novel fusion method based on convolutional neural network for malware detection. The model combines four classical convolutional neural network structures and its experimental results show that this method can effectively improve the accuracy and robustness of the model. The malware detection method based on convolutional neural network can avoid complex feature code extraction work, reduce labor costs, reduce system overhead, improve detection efficiency and detect new viruses. The training of the generative adversarial network in our model is difficult. In the future, we would like to research a more structured generation adversarial network training model to generate higher quality data.

REFERENCES

[1] Sami A, Yadegari B, Rahimi H, et al. Malware detection based on mining API calls: ACM Symposium on Applied Computing, 2010[C].

[2] M. Christodorescu, S. Jha. Static analysis of executables to detect malicious patterns[C].//In Proceedings of the 12th USENIX Security Symposium, 2003:169–186.

[3] A. Sung, J. Xu, P. Chavez, S. Mukkamala. Static analyzer of vicious executables (save)[C].//In Proceedings of the 2004 Annual Computer Security Applications Conference (ACSAC), 2004:326–334.

[4] R. Lo, K. Levitt, R. Olsson. MCF: a malicious code filter[C].\\Computers and Security 14,1995:541–566.

[5] SANTOS I,BREZO F,UGARTE-PEDRERO X,et al.Opcode sequences as representation of executables for data-mining-based unknown malware detection[J]. Information Sciences,2013,231(2013):62-82.

[6] NAKAZATO J,SONG J,ETO M. A novel malware clustering method using frequency of function call traces in parallel threads[J]. IEICE transactions on information and systems,2011,E94-D(11):2150-2158.

[7] Sundarkumar G G, Ravi V. Malware detection by text and data mining[C]. IEEE International Conference on Computational Intelligence and Computing Research. IEEE, 2013:1-6.

[8] Matthew G.Schultz,Eleazar Eskin,Erez Zadok.Data Mining Methods for Detection of New Malicious Executables [C].IEEE Computer Society,2001:38-49.

[9] Sundarkumar G G, Ravi V, Nwogu I, et.al. Malware detection via API calls, topic et al. Malware Detection Systems Baesed on API Log Data Mining[C]. IEEE, Computer Software and Applications Conference. IEEE Computer Society, 2015:255-260.

[10] Fujino A, Murakami J, Mori T, Discovering similar malware samples using API call topics[C]. Consumer Communications and NETWORKING Conference. IEEE, 2015:140-147.

[11] Hyrum S. Anderson, Phil Roth, et al. EMBER: An Open Dataset for Training Static PE Malware Machine Learning Models[C] Neuroscience School Of Advanced Studies(NSAS), 2018, 4.

[12] DAHL G E, STOKES J W, DENG L, et al. Large-scale malware classification using random projections and neural networks[C] 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Vancouver, BC, Canada: IEEE, 2013: 3422-3426.

[13] SAXE J, BERLIN K. Deep neural network based malware detection using two dimensional binary program features[C] 201510th International Conference on Malicious and Unwanted Software (MALWARE). Fajardo, Puerto Rico: IEEE, 2015: 11-20.

[14] KOLOSNJAJI B, ZARRAS A, WEBSTER G, et al. Deep learning for classification of malware system call sequences[C] Australasian Joint Conference on Artificial Intelligence. Hobart, TAS, Australia: Springer International Publishing, 2016: 137-149.

[15] TOBIYAMA S, YAMAGUCHI Y, SHIMADA H, et al. Malware detection with deep neural network using process behavior[C] 201640th Annual IEEE Conference on Computer Software and Applications (COMPSAC). Atlanta, GA, USA: IEEE, 2016, 2: 577-582.

[16] PASCANU R, STOKES J W, SANOSSIAN H, et al. Malware classification with recurrent networks[C] 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Brisbane, QLD, Australia: IEEE, 2015: 1916-1920.

[17] Edward Raff, Jon Barker, Jared Sylvester, Robert Brandon, Bryan Catanzaro, Charles Nicholas, et al. Malware Detection by Eating a Whole EXE[C], Neuroscience School Of Advanced Studies(NSAS). 2017, 10.

[18] Goodfellow, I., Bengio, Y., Courville, A.．Deep learning (Vol. 1)．Cambridge：MIT press， 2016.

[19] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, L., Wang, G.and Cai, J., 2015. Recent advances in convolutional neural networks. arXiv preprint arXiv:1512.07108.

[20] LeCun, Y. and Bengio, Y., 1995. Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks, 3361(10), 1995.

[21] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, L., Wang, G.and Cai, J., 2015. Recent advances in convolutional neural networks. arXiv preprint arXiv:1512.07108.

[22] Khan R U, Zhang X, Kumar R. Analysis of ResNet and GoogleNet models for malware detection[J]. Journal of Computer Virology and Hacking Techniques, 2019, 15(1): 29-37.

[23] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[24] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]. Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, 2012 : 1097-1105.

[25] SIMONYAN K, ZISSERMAN A J C. Very Deep Convolutional Networks for Large-Scale Image Recognition[J], 2015, abs/1409.1556.

[26] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[J], 2015, : 1-9.

[27] IOFFE S, SZEGEDY C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[C]. ICML, 2015 : .

[28] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the Inception Architecture for Computer Vision[J], 2016, : 2818-2826.

[29] SZEGEDY C, IOFFE S, VANHOUCKE V. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning[C]. AAAI, 2016 : .

[30] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[J], 2016, : 770- 778.

[31] WANG M, LIU B, FOROOSH H J I I C O C V W. Factorized Convolutional Neural Networks[J], 2017, : 545-553.

[32] RADFORD A, METZ L, CHINTALA S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks[J]. CoRR, 2015, abs/1511.06434.

[33] NATARAJ L, KARTHIKEYAN S, JACOB G, et al. Malware images: visualization and automatic classification[C]. Proceedings of the 8th International Symposium on Visualization for Cyber Security, 2011 : 1-7.

[34] Rathore H, Agarwal S, Sahay S K, et al. Malware Detection Using Machine Learning and Deep Learning[C]//International Conference on Big Data Analytics. Springer, Cham, 2018: 402-411.

[35] Sewak M, Sahay S K, Rathore H. Comparison of deep learning and the classical machine learning algorithm for the malware detection[C]//2018 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD). IEEE, 2018: 293-296.

[36] Gibert D, Mateu C, Planes J. The rise of machine learning for detection and classification of malware: Research developments, trends and challenges[J]. Journal of Network and Computer Applications, 2020: 102526.

AUTHORS

I am a postgraduate student of the Institute of Information Engineering, University of Chinese Academy of Sciences, Majoring in deep learning and security.

TRUSTED COMPUTING IN DATA SCIENCE: VIABLE COUNTERMEASURE IN RISK MANAGEMENT PLAN

Uchechukwu Emejeamara[1], Udochukwu Nwoduh[2] and Andrew Madu[2]

[1]IEEE Computer Society, Connecticut Section, USA
[2]Department of Computer Science, Federal Polytechnic Nekede, Nigeria.

ABSTRACT

The need for secure data systems has prompted, the constant reinforcement of security systems in the attempt to prevent and mitigate risks associated with information security. The purpose of this paper is to examine the effectiveness of trusted computing in data science as a countermeasure in risk management planning. In the information age, it is evident that companies cannot ignore the impact of data, specifically big data, in the decision making processes. It promotes not only the proactive capacity to prevent unwarranted situations while exploiting opportunities but also the keeping up of the pace of market competition. However, since the overreliance on data exposes the company, trusted computing components are necessary to guarantee that data acquired, stored, and processed remains secure from internal and external malice. Numerous measures can be adopted to counter the risks associated with data exploitation and exposure due to data science practices. Nonetheless, trusted computing is a reasonable point to begin with, in the aim to protect provenance systems and big data systems through the establishment of a 'chain of trust' among the various computing components and platforms. The research reveals that trusted computing is most effective when combined with other hardware-based security solutions since attack vectors can follow diverse paths. The results demonstrate the potential that the technology provides for application in risk management.

KEYWORDS

Trusted Computing, Security, Data, Data Science, Provenance, Risk Management, Big Data, Trusted Platform Module, Platform Computation Register

1. INTRODUCTION

The constantly growing need for information has prompted data scientists to create more advanced models that can improve the gathering and analysis of big data. However, as stated by FOX [10], most technological developments come along with numerous risks that impact significantly on the information systems of companies. As a result, not only may the companies fail to achieve the intended results but also risk losing significant information including its big data. Faced with these increasing threats, cyber security teams and data scientists have sought refuge in trusted computing that creates a 'chain of trust' among the various computing platforms. The trusted computing block is enforced with numerous chained signatures and encryptions that attempt to prevent the successful malice that would appear to be successful in their absence [5]. Since data is such a critical resource, companies lack the freedom to expose it recklessly to agents of malice due to the costs that would be incurred. These costs can come in the form of customer loyalty, lawsuits, loss of competitive edge, and loss of revenue among

David C. Wyld et al. (Eds): SPTM, IPPR, AIS, CSIT, DaMi, AMLA - 2020
pp. 17-26, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100602

others. The costs are adequate to hamper the growth of the company. This paper proposes a feasible risk management plan that utilizes trusted computing technologies to protect the company information during the utility of data through various data science approaches. This paper also reveals pertinent features of TC that makes it a viable countermeasure for inclusion in Risk Management Plans for organizations. It helps in the generation of cryptographic keys that prevent the modification of the software or the software data where it is applied. TC can be used as a component of other recognized security countermeasures, including authentication, IDS, firewalls, access controls, and VPN through software modifications and enhancements to the hardware.

2. TECHNICAL SURVEY

2.1. Trusted Provenance

Data integrity is a primary goal for information security systems. To achieve this objective, one of the primary approaches is to guarantee the reliability of provenance systems. Provenance systems are centered on the concept that the original data used in the formation of certain systems should be safeguarded for auditing and reference. When faced with security challenges, system auditors can always retreat to the original framework that created a system with code and infrastructure to understand the actual source of the problem. This concept brings to light the necessity of trusted computing in provenance systems. Since provenance systems have such value when addressing security challenges, Lyle and Martin [6] explain that failure to have 'trusted pervasive hardware infrastructure' would only lead to increased susceptibility of the information systems to attackers. Lyle and Martin [6] adds that malicious agents, target the provenance system since they can best inform them of the details of the specific infrastructural system. Hence, maintaining secure provenance systems is critical towards the protection of the data systems within any organization.

Companies are over-relying on external data for their information. This trend is evident from the unmatched utility of big data and analytics tools in the management's decision-making engagements. However, cyber threats evolve with evolving technologies and business needs. In this way, cyber-attacks target company information either in its storage form, which Bao, Chen, and Obaidat [9] refer to as 'data at rest', data being processed, or data at transit. While these threats are evident, data scientists and security teams have the task to ensure that their data is safe from attacks. One of the simplest ways for attackers to successfully infiltrate information systems, according to Hu et al. [4], is to gain access to provenance information. This information, as explained by Hu et al. [4], gives hackers the root resources about the target system paralleling their knowledge of their target system with that of the system owners. Problematically, the stakes that hackers with provenance information will not just steal information but overrun a system or obliterate its resources, are rather high [6]. In this respect, the risk factors for data loss or obliteration when provenance information is stolen are significantly critical that the best or possibly the only solution would be to prevent successful infiltration.

2.2. Security Risks in Data Science

Security in data science can take various forms. The basic security approaches include security on the software resources, hardware or infrastructural resources, data protection as a security concept independent from the former two, and data anonymization. Gordo [2] explains that security in data science would also include information warfare due to the increasingly high rates of availability of strategically deceptive data meant to promote misinformed moves by various corporations and institutions. It has taken new approaches that include significant turns of events since the birth of psychologically deceptive approaches such as social engineering. In this context, information warfare has been one of the most significant challenges in data science leading data scientists, to get misinformed patterns about the data available. Moreover, the correctness of data can also render it obsolete when its usability is compromised by hackers. Hence, having the right information must always be paired with reliably secure systems to make viable use of any type of data.

The challenges and solutions to data security revolve around the physical infrastructure, software infrastructure, the data, and people within and around an organization. Physical infrastructure security protects data from technical failures and physical and virtual malice agents. While data scientists have the task to remain vigilant, infrastructure resources are managed by the information security teams. The security of the infrastructure is a primary factor to guarantee the security of the data within its systems. Software infrastructure focuses on the vulnerabilities that could be exploited by hackers. It supplements hardware infrastructure by creating frameworks using which virtual attacks can be detected, prevented, and repelled. In other cases, the software can take the usual task of informing the computer security incident response team to include human decisions in the process. Data security is another issue outside the ordinary context of information security fostered by infrastructural components. However, data scientists with security skills can identify anomalies in data that alert them of falsifications [3]. Such trends can protect data scientists from the risk of using false information that should also alert them of a security challenge in their information systems. Moreover, encryption promotes data anonymity regardless of the security standards applied by an organization. Finally, the people in any organization determine significantly the security of data within the data systems. Essentially, the organizational culture based on its dedication to follow the set policies explains the accountability to protect data resources and prevent insider threats. Nonetheless, all these components of information security must be effective to establish the chain of trust among them and achieve the necessary security standards.

2.3. Specific Data Science Challenges

Privacy and security remain the topmost challenges in data science. Cai and Zhu [7] explains that the bigger the data a company holds, the more susceptible it is to attackers. This susceptibility raises the question of the privacy of information stored within the information systems. Notably, while confidential information often revolves around customer information, data science confidentiality may take a slightly different perspective and focus on the quality of information achieved from interpreting the data. Most decisions are made based on the information available which they would like to conceal. Notably, this information also forms part of their competitive advantage, which, if adequately sustainable, would lead other companies to target competitors with better selling propositions for the market. Regardless of the type of data used, privacy must be guaranteed to prevent lawsuits which can lead to significant losses in revenue and tainting of the brand image.

The increasing significance of data in corporate decision-making processes has made the use of data compulsory. According to Ingrams [1], big data utility is not only fostered by the need for competitive advantage but also the opportunity costs of ignoring its utility. In explanation, improved security has been achieved through the analysis of big data that provides proactive insights into the security issues that could be affecting other companies and could happen at a specific company. Ignoring such aspects keeps the company away from information leading them to make uninformed security decisions that could risk the data of the company. However, while this opportunity cost is irresponsible for any security teams to incur, data science practices such as data collection, evaluation of its validity, data analysis, and decision making are, by itself, a complex and costly engagement. The resources required are overly expensive, and the information may not necessarily be useful as to cover the costs if the process is not well-informed. For such reasons, Data Scientists employed must be well trained and experienced professionals, to keep cost within considerable limits.

By contrast, increasing challenges in the area of data science have been caused by the nature of data used, the tools, or the skills. Essentially, before data scientists can assume that the data will give the right results, they need to ensure the data itself is the right one. Fig 1, highlights the danger posed by disorganized, inaccurate, and incoherent data. Bad data not only costs the company the resources consumed but also the misinformed decisions that will be made. Moreover, the company is more likely to fall behind competition due to the failure to draw correct business insights [7]. Bad analytics, on the other hand, results in the misinterpretation of data. Bad analytics can take the form of using the wrong tools or using the right tools without the foreknowledge of the best data science practices. The consequence is the derivation of wrong patterns or their misinterpretation. Similarly, the risk, in this case, is still the lagging behind the competition.

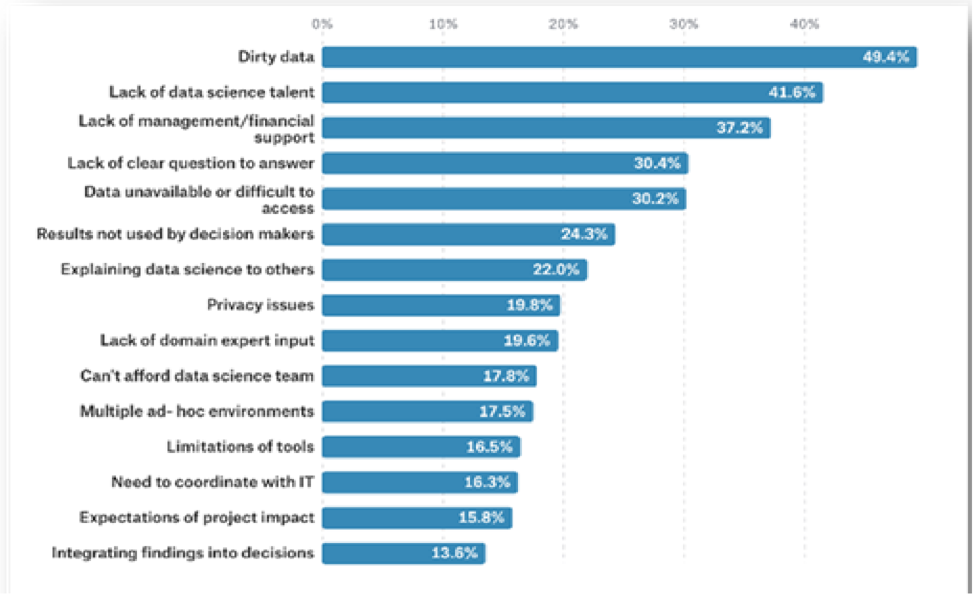


Fig. 1 Significance of Challenges in Data Science [11]

3. METHODS

The Risk Management Plan (RMP) will be implemented similarly to any other security approach. As trusted computing pertains more to the individual components, scholarly works will be used to assess the effectiveness of the approaches.

3.1. Risk Identification

The first process in risk planning is the identification of the risk aspect. The risk factor associated with data science with regards to computing has been identified by numerous scholars to be highly significant [7]. The infrastructural, data, and people-related security aspects, for instance, would have a considerable impact on the risk positioning of the company when considered. If ignored, data science practices could be rendered not only obsolete but also expensive. Additionally, issues with big data pose a significant risk since they can foster bad decisions when analyzed incorrectly or when the data is not right. Such instances can pose a dangerous cost risk should the misinformed decisions affect crucial operations of the company. Moreover, according to Bilić [8], data science is at such risks from the falsification of information by intentional agents to mislead their competition into arriving at the wrong conclusions. Hence, the identification of these risks is the first step in creating sufficient countermeasures in trusted computing with regards to data science.

3.2. Risk Assessment

The assessment of a risk factor helps teams to substantiate its threat and determine the weight of the associated risk. Three common risk factors are assessed that include threats, malware, and anomalies. In the context of data science, some aspects of assessing anomalies use trusted computing principles that seek to find consistency between data sets such as information stored in the PCRs [6]. By contrast, anomalies can arise from the data itself due to falsifications or decreased security. When the software components identify such anomalies, trusted computing is compromised since the chain of trust no longer exists. Similarly, malware risks can be assessed to confirm the risk factors they pose. Notably, while machine learning has been improving the utility of big data, cyber threats have increased as they attempt to use the same principles such as dynamic code analysis to compromise the effectiveness of the data science machine learning algorithms. The risk, in this case is so significant that it can threaten the company's sustainability as it lags behind competition due to ill-informed decisions. Trusted computing also aims at identifying threats, such as logs of failed and successful authentications and be able to explain the incidents. The assessment of such incidents can explain how targeted the company is and the risk posed by any threats.

3.3. Risk Control

The final process in the risk management plan is the control of the identified risks. The risks are diverse and numerous approaches must be applied to fully address the challenges. Provenance systems can expose the company considerably. The chain of trust between the computing systems in the data science computing environment would require subtle implementations, such as the use of endorsement keys stored in the TPM, memory curtaining that isolates critical memory components from the less sensitive ones, sealed storage that binds information to its infrastructural components, a remote attestation that connects with legitimate parties remotely, and trusted third party that is based on the security measures of a remote computing base [5]. However, other than having these security solutions in place, ethical data scientists with adequate skills and resources are also at the bottom line of controlling the risks, otherwise, the costs would be dire. Overall, the risk control process is a collaborative effort between all stakeholders to guarantee the achievement of the desired outcomes. Fig. 2 depicts the process flow in RMP.

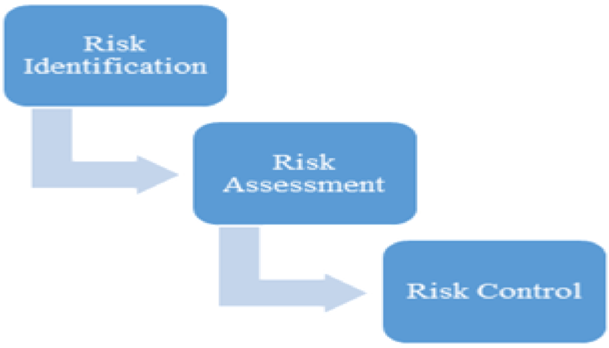


Fig. 2 RMP Flow

3.4. Trusted Computing

Trusted Computing (TC) is a hardware-based method that enforces the integrity of software or platforms. It protects a system from software-level attacks [19]. It operates on two principle ideas, namely remote attestation and sealed memory [12]. Remote attestation is concerned with the identification and declaration of the software that runs on a remote computer while sealed memory authorizes specific software stacks to access stored secrets. Maene et al. [12] explained that TC provides both local and remote attestation. Through sealing, TC wraps data in a manner that prevents unwrapping without the decrypting key. From these two foundational ideas, TC enables the development of security protocols, including authentication, encryption, and the management of digital rights. The simple framework for achieving trust using TC is as depicted in Fig. 3.

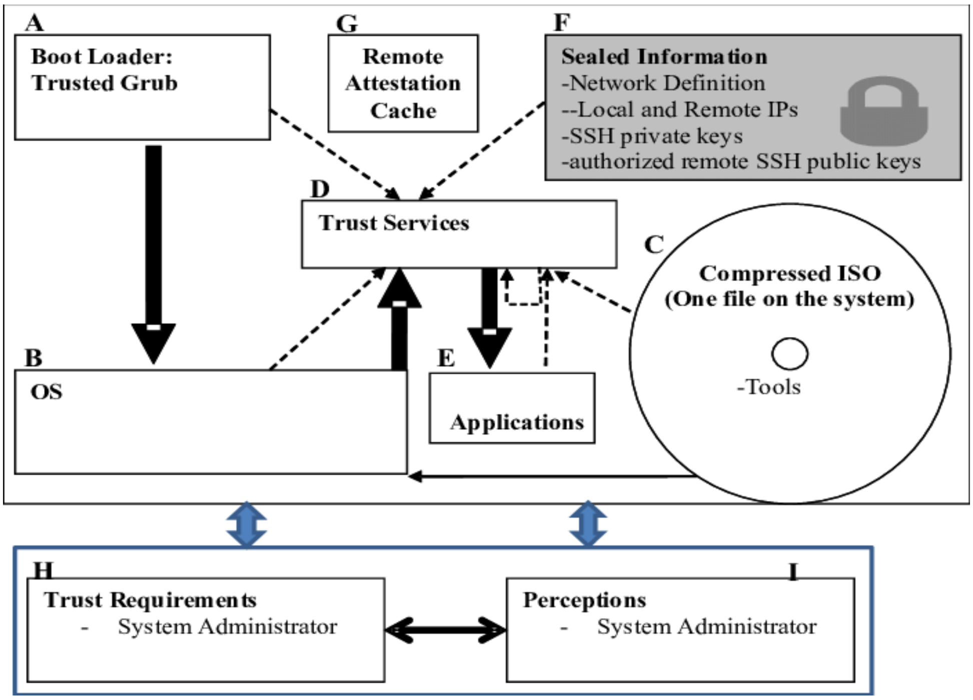


Fig. 3 Trust development framework based on TC

With the proliferation of mobile devices that rely heavily on digital data, TC provides an effective scheme for securing the data in applications installed on various smartphones. Using a Trusted Execution Environment (TEE), Fan et al. [13] applied the TC framework to minimize data losses from mobile devices. The TEE enabled the authors to divide sensitive files into file slices, encrypt the individual file slices/parts, and map individual file parts to whole files in cloud computing [13]. They further reported that the scheme thus developed was highly efficient and secure. According to Maene et al. [12], the trusted computing architecture guarantees users the protection against software-level attackers, thereby acting as a countermeasure against the exploitation of software vulnerabilities. The most practical use of TC is found when it is integrated with other security protocols.

4. DISCUSSION

Risk management teams have no option but to include IT security comprehensively into their plans. The risk factor with data science is rather critical that the failure to promote trusted computing among data-handling components can render the data science operations obsolete. In explanation, trustworthy behaviour is enforced through the use of standards that ensure a ‘chain of trust’ between the hardware and software components that have security roles to play. The Trusted Platform Module (TPM), for instance, stores isolated encryption keys that are used by the users to authenticate various processes and platform computation registers which hold integrity registers [6]. The TPM, essentially, guarantees that records are maintained regarding the various computing processes that are necessary for auditing to confirm that the normal processes have not been compromised. However, trusted computing uses backup values about processes to validate the processes such as that any inconsistencies between the data would require an explanation. Hence, trusted computing can ensure that anomalies are addressed when components do not function as expected.

4.1. Efficacy of Trusted Computing

Trusted Computing, in the context of data science, can be used as a countermeasure in Risk Management. Trusted Computing (TC) is effective in securing data, making it an instrumental technology in data science. In mobile technologies and cloud computing, Trusted Computing enables the minimization of data losses through file slicing and slice encryption and allows for the mapping of file slices to cloud computing environments [13]. The technology is also effective in securing applications that reside on mobile devices from unauthorized access. Trusted Computing also provides an effective solution for promoting the privacy and security of data stored in cloud infrastructure through the use of technologies such as Intel Software Guard Extensions (SGX) [15]. Thus, organizations that provide Software-as-a-Service (SaaS) cloud solutions can effectively implement TC-based privacy and security policies. When used along with hardware-based solutions such as TEE and other security systems, Trusted Computing becomes an effective security solution. Its effectiveness has been proved for protection against unauthorized access, man-in-the-middle attacks, and password guessing attacks [14].

The practicable solutions that TC offers, based on its integration with other security technologies, makes it useful in RMP countermeasures in the contexts of data science, cloud computing, Internet of Things (IoT), and blockchain applications. In IoT solutions, TC provides a platform for addressing the challenge of users’ loss of control over data, as evidenced by the HyperNet framework [16]. In blockchain applications, Trusted Computing can be used to promote user control over their data through the Proof-of-Credibility option that it offers [17]. Furthermore, Trusted Computing principles allow for the development of various trust enhancement frameworks that are a necessity in blockchain applications [18]. Thus, Trusted Computing finds applications in technologies that inform the future of the data science landscape.

While it protects specific applications from attacks, Trusted Computing is not an effective countermeasure against attacks that target other sections of a computing system. The attacker model it adopts assumes the ability of the cybercriminal to tamper with the OS, launch malicious software, sniff the network, perform MITM attacks, modify traffic, break network-based cryptographic primitives, and launch denial-of-service (DoS) attacks [12]. Consequently, Trusted Computing is not an effective countermeasure against the threats facing an information system unless it is used along with other hardware-based security solutions that protect the OS, network, and traffic.

4.2. Trusted Computing Viability in Risk Management

From literary research, trusted computing is a significant aspect of IT security. It promotes the availability of mutually secure systems due to the increased security in each specific component. One of the most significant benefits of trusted computing is the capacity to protect data systems through the presence of built-in processes. These processes revolve around encryptions and hashing preventing successful attacks. Notably, data encryption and hashing is an added layer of security that ensures that information remains confidential even in the event of successful infiltration. This way, the data is protected from compromise.

Being an approach for establishing trust between components, trusted computing also guarantees that other systems are secure. In explanation, trust can only be established between the components when each of them is secure from malice. In this way, the approach seems to play the same role in promoting the security of all infrastructural components in a distributed computing environment. This aspect is a limitation since it seems to give trusted computing a broad connotation as an umbrella term, which it is not. Moreover, trusted computing acts as a facilitator for creating a safer environment even when resources are decentralized from the main systems.

4.3. Challenges

Trusted computing works at a rather low level of computing. Although it is one of the best ways of protecting data systems, trusted computing on its own may not be able to guarantee that security will be achieved. For instance, trusted computing can perform the best task to promote the security of provenance systems. However, attack vectors do not necessarily follow that path. Other approaches such as social engineering have been used extensively that organizations can only focus on trusted computing alone as the viable countermeasure to mitigate the risks associated with threats in data science.

4.4. Verdict

The presence of a plethora of vectors posits that companies cannot focus on one of them. However, the significance of any of them should not be overlooked since they add up to the overall security of the data systems within any organization. Trusted computing offers significant advantages for its users concerning data science. With the basic security challenges having been solved, the additional approaches can guarantee a comprehensive and adequate addressing the security concerns that revolve around the use of data within and external to a company. For this reason, trusted computing is a viable countermeasure in risk management planning but can only be applied as a complementary strategy reinforcing other security mechanisms.

5. CONCLUSION

Trusted computing is one of the primary approaches that use security methods outside the common software-related mechanisms. The understanding of the significance of trust in security begins with the foreknowledge of the risk impacts of not improving the effectiveness of the various security and protection of information systems. Trust is established when all components can guarantee the security that the data being stored, processed, or in transit in these systems is secure from both internal and external threats. With this security guaranteed, the remaining risk mitigation measure would be the assurance that the data scientists perform the right tasks correctly and with the right tools. In this way, the countermeasures for the risk management plan would be viable.

REFERENCES

[1] A. Ingrams, "Public Values in the Age of Big Data: A Public Information Perspective", *Policy & Internet*, vol. 11, no. 2, pp. 128-148, 2018. Available: 10.1002/poi3.193.

[2] B. Gordo, "“Big Data” in the Information Age", *City & Community*, vol. 16, no. 1, pp. 16-19, 2017. Available: 10.1111/cico.12219.

[3] B. Tellenbach, M. Rennhard and R. Schweizer, "Security of Data Science and Data Science for Security", *Applied Data Science*, pp. 265-288, 2019. Available: 10.1007/978-3-030-11821-1_15.

[4] D. Hu, D. Feng, Y. Xie, G. Xu, X. Gu and D. Long, "Efficient Provenance Management via Clustering and Hybrid Storage in Big Data Environments", *IEEE Transactions on Big Data*, pp. 1-1, 2019. Available: 10.1109/tbdata.2019.2907116.

[5] E. Padma, "Trusted Attestation System for Cloud Computing Environment Using Trusted Platform Module", *Internet of Things and Cloud Computing*, vol. 5, no. 3, p. 38, 2017. Available: 10.11648/j.iotcc.20170503.11.

[6] J. Lyle and A. Martin, *Trusted Computing and Provenance: Better Together*. Oxford: Oxford University Computing Laboratory, 2010.

[7] L. Cai and Y. Zhu, "The Challenges of Data Quality and Data Quality Assessment in the Big Data Era", *Data Science Journal*, vol. 14, no. 0, p. 2, 2015. Available: 10.5334/dsj-2015-002.

[8] P. Bilić, "Search algorithms, hidden labour and information control", *Big Data & Society*, vol. 3, no. 1, p. 205395171665215, 2016. Available: 10.1177/2053951716652159.

[9] R. Bao, Z. Chen and M. Obaidat, "Challenges and techniques in Big data security and privacy: A review", *Security and Privacy*, vol. 1, no. 4, p. e13, 2018. Available: 10.1002/spy2.13.

[10] S. FOX, "Policing - The technological revolution: Opportunities & challenges!", *Technology in Society*, vol. 56, pp. 69-78, 2019. Available: 10.1016/j.techsoc.2018.09.006.

[11] N. Thabet and T. Soomro, "Big Data Challenges", *Journal on Computer Engineering and Information Technology*, vol. 4, no. 3, 2015. Available: 10.4172/2324-9307.1000133.

[12] P. Maene, J. Götzfried, R. d. Clercq, T. Müller, F. Freiling and I. Verbauwhede, "Hardware-based trusted computing architectures for isolation and attestation," *IEEE Transactions on Computers*, vol. 67, no. 3, pp. 361-374, 2018.

[13] Y. Fan, S. Liu, G. Tan, X. Lin, G. Zhao and J. Bai, "One secure access scheme based on trusted execution environment," in *12th IEEE International Conference On Big Data Science And Engineering*, New York, NY, 2018.

[14] E. F. Cahyadi, Y.-C. Chou, C.-Y. Yang and M.-S. Hwang, "An improved mutual authentication scheme with smart cards and password under trusted computing," in *2018 the 2nd annual International Conference on Cloud Technology and Communication Engineering*, Nanjing, China, 2018.

[15] A. T. Gjerdrum, R. Pettersen, H. D. Johansen and D. Johansen, "Performance principles for trusted computing with Intel SGX," in *International Conference on Cloud Computing and Services Science*, Porto, Portugal, 2017.

[16] H. Yin, D. Guo, K. Wang, Z. Jiang, Y. Lyu and J. Xing, "Hyperconnected network: A decentralized trusted computing and networking paradigm," *IEEE Network*, vol. 32, no. 1, pp. 112-117, 2018.

[17] D. Fu and L. Fang, "Blockchain-based trusted computing in social network," in *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, Chengdu, China, 2016.

[18] Y. Wu, Y. Qiao, Y. Ye and B. Lee, "Towards improved trust in threat intelligence sharing using blockchain and trusted computing," in *2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*, Granada, Spain, 2019.

[19] M. Alhaidary, S. M. M. Rahman, M. Zakariah, M. S. Hossain and A. Alamri, "Vulnerability analysis for the authentication protocols in trusted computing platforms and a proposed enhancement of the OffPAD protocol," *IEEE Access*, vol. 6, pp. 6071-6081, 2018.

GRANT-FREE SAFE-SCMA BASED ON DETECTION OF UNKNOWN ABNORMAL CODEBOOKS

Hanyuan Huang, Tao Li and Hui Zhao

College of Cyber security, Sichuan University, Chengdu, China

ABSTRACT

Non-orthogonal multiple access (NOMA) can support massive accesses in 5G. Sparse code multiple access (SCMA) is a typical NOMA technology. Basic principle of SCMA is that multi-user bit data directly map into multi-dimensional complex sequences through codebooks. Grant-free SCMA allows users to select codebooks from codebook pool to send data instantly, reducing the cost of overhead and delay of granting process. When the receiver and the sender use same codebook information, the data can be transmitted correctly. But in current SCMA researches, the problem of asymmetric codebook information between sender and receiver caused by the intrusion of codebook pool is not considered. In this paper, abnormal codebook detection is proposed in the grant-free SCMA. Because most of intrusion is unknown, initial detection is realised by comparing characteristics of tested codebooks with those legal ones concluded during codebook design process. In this paper, tested objects in the test process can include but not limit to codebook structure, constellations, distribution of constellations, overall feature of codebook pool. Test is executed until discovering error states or accomplishing all tested contents. Inspired by the distinction between self and non self in the artificial immune system, tested abnormal codebooks are saved to act as detectors. To take full advantages of known non-self codebooks, saved detectors are further evolved, and future detection can do match with detectors which are evolved from those known abnormal codebooks to discover some kinds of unknown abnormal codebooks.

KEYWORDS

Grant-free SCMA, SCMA codebooks, abnormal codebooks detection, artificial immune system, abnormal codebook evolution

1. INTRODUCTION

The requirements in 5G include high spectrum efficiency, low latency and massive access. Non-orthogonal multiple access (NOMA) [1-3] allows massive users to share same resources to satisfy these requirements. NOMA is an alternative access technique in 5G. And Sparse Code Multiple Access (SCMA) is a typical example of NOMA. It is able to be demodulated with reasonable complexity by message passing algorithm (MPA) [4]. The principle of SCMA is that multi-user binary data directly maps to multidimensional sparse codewords of SCMA codebooks, where each codeword represents a spread transmission layer, and different layers transmit on the same time-frequency resources. The feature of SCMA multi-user codebooks is one of the keys deciding system performance [5]. The codebooks used for demodulation by the receiver is exactly the same as those used for modulation by the transmitter, which is the necessary prerequisite for correct transmission. In order to save the extra overhead and latency with granting, grant-free SCMA allows users to do data transmission instantly. But the receiver does

David C. Wyld et al. (Eds): SPTM, IPPR, AIS, CSIT, DaMi, AMLA - 2020
pp. 27-33, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100603

not know about which users are currently transmitting data. Because of the instant transmission, all users are less likely to send data at the same time. If each user is assigned a fixed codebook, it will reduce utilization of codebooks. Therefore, in grant-free SCMA, the active users who need to send data select codebooks randomly from the codebook pool for data modulation. The receiver determines the codebook information of active users through corresponding relationship between pilots and codebooks [6].

When problems about security of codebook pool occur, there will be some abnormal codebooks in the codebook pool that the receiver does not contain. If active users select abnormal codebooks for data modulation, receiver will not be able to complete the data demodulation correctly. Therefore, this paper proposes an abnormal codebook detection mechanism. After a user selects a codebook and before modulation, the validity of the codebook is tested firstly. One of the major algorithms to discriminate self and nonself within field of artificial immune systems (AIS) [7] is Negative Selection Algorithm [8]. In NSA, initial detectors are generated randomly and practised by self tolerance to turn into mature detectors. If data matches a detector, it can be seen as abnormal. Because the system is unknown for most abnormal codebooks, the known legal codebook design is prior information in the detection process. A codebook test and detector generation method based on codebook features is proposed, avoiding random detector generation and self tolerance with low efficiency. The test contents can include but not limit to codebook structure, codebook constellations, correlation between constellations. When error states are detected or all contents are completed, the current detection is stopped. If error states are detected, the tested codebook is determined to be abnormal codebook and added to the detector set. To cover the shortage that information of known non-self ones are not taken full advantages of in the conventional NSA, on the premise that a codebook which is very similar to the abnormal codebook can be seen as an abnormal one, the malicious codebook is evolved and stored as detector. The similarity concludes that close Euclidean distance between constellations, or the codebook structure is similar and so on. When the user selects a codebook next time, a tested codebook can match with detectors firstly. If the matching is successful, the codebook will be determined as an abnormal codebook. If the number of detectors does not meet the requirements, codebook test based on codebook features and detector generation are executed continuously.

2. SYSTEM MODELS

2.1. SCMA Modulation

A SCMA [9] encoder can be seen as a mapper from $\log_2 M$ bits to a K-dimensional sparse codeword of a codebook and the size of the codebook is $K \times M$. There are Z ($Z < K$) non-zeros elements in a codeword. Each codeword represents a spread transmission layer. There are J layers, where the j-th layer is expressed as $\mathbf{x}_j=[x_j(1), x_j(2),\ldots, x_j(K)]^T$, and $\mathrm{E}\left[\left\|\mathbf{x}_j(k)\right\|^2\right]=1$. The received signal vector $\mathbf{y}$ over all the subcarriers can be expressed as:

$$\mathbf{y}=\sum_{j=1}^{J} \mathrm{diag}\big(\mathbf{h}_j\big)\mathbf{x}_j+\mathbf{n} \tag{1}$$

Where $\mathbf{h}_j=[h_j(1),h_j(2),\ldots,h_j(K)]^T$ is the channel gain for the j-th layer. And diag(·) refers to a diagonal matrix, in which the k-th element on the diagonal is $h_j(k)$. The process of SCMA is shown in Figure.1 with M=4, K=4 and Z=2.

The received signal in the k-th subcarrier resource is:

$$y(k) = \sum_{j\in\varphi_k} h_j(k)x_j(k) + n(k) \tag{2}$$

Where $h_j(k)$ is the channel gain for the j-th layer in the k-th resource and $n(k)$ is complex Gaussian noise in the k-th resource. φ_k is a set made up with all the layers in the k-th resource and its cardinality is d_f, which means d_f layers occupying the k-th resource.

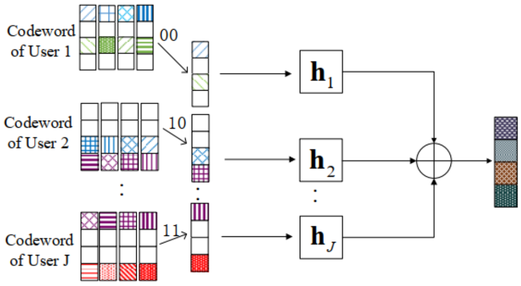


Figure 1. The process of SCMA (J = 6, K = 4, Z = 2).

2.2. SCMA Demodulation

Message Passing Algorithm (MPA) is usually used to do SCMA decoding. MPA is an iterative detection algorithm with optimal BER (Bit Error Rate) performance and acceptable computational complexity. MPA estimates the transmitted vector $\tilde{\mathbf{x}}_j$through the received signal $\mathbf{y}$. The initialization of confidential information in the k-th resource of MPA can be written as:

$$\Phi_k = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\left\|y(k) - \sum_{j\in\varphi_k} h_j(k)x_j(k)\right\|^2\right) \tag{3}$$

Where σ refers to the noise variance. Each MPA iteration consists of two steps: i) passing confidential information from resource nodes to user nodes; ii) information exchange from user nodes to resource nodes. Then, a posteriori probability for the corresponding codeword $\mathbf{x}_j$. The binary log-likelihood ratios (LLRs) is used to decide what the $\log_2 M$ bits are.

The basic condition for correct SCMA transmission is keeping equal codebook information between transmitter and receiver.

2.3. Grant-free SCMA

In grant-free transmission, terminals will not wait for a transmission grant from the base station, which saves a lot of overhead and latency with granting. We defined users doing the transmission as active users, and others having no transmission as inactive users. When an active user needs to send message, it randomly selects a codebook from the codebook pool to modulate bit data. There are N codebooks totally. Active users select codebooks randomly from N codebooks to modulate data. The receiver can do active user detection and confirm codebooks being selected through active pilots detection based FOCUSS or EM algorithm [10]. It is possible that different active users select same codebook. But the receiver still can demodulate data with inferior but not nonlethal performance, as long as information of codebook pool saved by receiver is same as that in transmitter.

3. SCMA CODEBOOKS

The SCMA codebook structure can be represented by factor graph or mapping matrix. The edge between layer node and resource node in factor graph or the non-zero element in k-th row and j-th column of mapping matrix means that the k-th resource is occupied by the j-th layer. Mapping matrix of a simple codebook example, where M=4, J=6, K=4, d_f=3, Z=2 in [11] is shown as:

$$\mathrm{F}_{4\times 6}=\begin{bmatrix}1&0&1&0&1&0\\0&1&1&0&0&1\\1&0&0&1&0&1\\0&1&0&1&1&0\end{bmatrix} \tag{4}$$

A layer is equipped with a complex constellations set S_u (u = 0, 1, · · · , $d_f - 1$) in one-dimensional resource. The cardinality of S_u is M. All elements in S_0 are reals and other $S_u = e^{j\theta_u}S_0$. The factor graph of the codebook example in [11] is shown in Figure.2 and $y(k)$, $h_j(k)$, $x_j(k)$ express same meanings as Eq. (2). The overload rate is λ = J/K = 150%. Besides, $x_1(1) \in S_0$, $x_3(1) \in S_1$, $x_5(1) \in S_2$. And constellations points in Fig.2 express that $S_1 = e^{j\theta_1}S_0$ and $S_2 = e^{j\theta_2}S_0$, $\theta_1 < \theta_2 \ldots < \theta_{d_f-1}$. Besides, the combination $\left\{S_0, \ldots, S_{d_f-1}\right\}$ is reused in every resource, which means an example of constellations allocation in all resources based on Eq.(4) can be expresses as:

$$\mathrm{F}_{4\times 6}=\begin{bmatrix}S_0&0&S_1&0&S_2&0\\0&S_2&S_0&0&0&S_1\\S_1&0&0&S_0&0&S_2\\0&S_0&0&S_2&S_1&0\end{bmatrix} \tag{5}$$

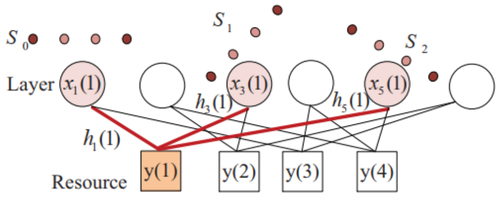


Figure 2. Factor graph of a SCMA codebook (J = 6, K = 4, d_f = 3, Z = 2).

In the iterative process of MPA demodulation, after calculating the confidential information of all codeword symbols on each resource separately, the confidential information is transferred between different resources and users. So, when considering the problems about constellation points, such as extraction of codebook feature and codebook design, it is feasible to be focus on those of one-dimensional resource. And when considering problems about codebook structure or constellations allocation in entire codebook, it is still necessary to think about all resources.

4. SAFE-SCMA DETECTING UNKNOWN ABNORMAL CODEBOOKS

In the condition that most of intrusion is unknown, useful prior information is normal features in the codebook design. After a user selects a codebook from codebook pool, the codebook should be estimated its validity. Obviously if normal codebooks are much more than abnormal codebooks, the probability that an abnormal one is selected is fairly low. It can satisfy high-reliability requirements in 5G. So, it is necessary to have a detection with large number of normal codebooks and acceptable complexity of computation. AIS has nature advantages in unknown intrusion detection because of self renewal and evolution. But discrimination between self and nonself in the typical algorithm NSA spends high computation complexity in random initial detector generation and tolerance with self set. It will have a strong impact on low-latency business in 5G. In the initial codebook test without detectors we propose a method based on codebook feature extraction. And tested abnormal codebooks are saved and evolved to act as detectors, which overcomes the shortcomings with aimlessness of detector generation and insufficient use of known abnormal data in the traditional NSA.

4.1. Initial Codebook Test and Detector Generation

Initial codebook test and detector generation based on codebook features are proposed, avoiding random generation and matching with all self codebooks. The test contents are based on the extraction of normal codebook features. The test process is carried out until the abnormal state is detected or all test steps are completed. A codebook with abnormal states can act as a detector. The following contents are about some codebook features for test.

4.1.1. Codebook Structure

There is a $K' \times M'$ tested codebook. It is set of K'-dimensional sparse codewords. There are Z' ($Z' < K'$) non-zeros elements in a codeword. And a normal codebook should be a $K \times M$one with Z non-zeros elements in a codeword. We can compare K' and K, M' and M, Z' and Z. If results are not all equivalence, the tested codebook is judged as an abnormal one. When comparison results are coincident, positions of non-zero elements in each codewords should be same and conform to one column of the mapping matrix (Eq. (4) is an example) in the codebook pool.

4.1.2. Values of Constellations

There is a $K' \times M'$ tested codebook. It is a set of K'-dimensional sparse codewords. Extract non-zero elements from one resource and note as $\{C_1, \dots, C_{M'}\}$. Firstly, detect whether these elements are on a straight line in the complex rectangular coordinate system. If so, then see whether the angle corresponding to this line belongs to $\left\{\theta_1, \dots, \theta_{d_f-1}\right\}$ decided in the codebook design. If the tested codebook cannot meet these conditions, it is an abnormal one.

4.1.3. Power of Codewords

The m-th codeword in the tested codebook is noted as $\mathbf{x}_m$=$[x_m(1), x_m(2),\dots, x_m(K')]^T$. To keep the fair of power allocation, different codewords should have the same average power, which means $\mathrm{E}\left[\|\mathbf{x}_m(k)\|^2\right]$=1.

4.1.4. Distribution of constellations

Non-zero elements on one resource form a straight line in the complex rectangular coordinate system. It can be found that Z' lines can be extracted in a tested codebook. And there are Z' angles $\{\theta'_1, \dots, \theta'_{Z'}\}$. We can determine constellation distribution of the tested codebook should belong to which column in Eq. (5). Because it is decided that $S_u = e^{j\theta_u} S_0$, $\theta_1 < \theta_2 \dots < \theta_{d_f-1}$, and differences between $\theta_u$in codebook design, we can judge whether the distribution of constellations meets the criterion in Eq. (5) through the difference between angles.

4.2. Detector Evolution

In conventional AIS, when a new detector generates, the process of random generation and tolerance with large self set are repeated. Previous detection information is not used for detector generation. It causes poor efficiency and waste of previous work. If an abnormal codebook that has been detected is evolved, the evolved results can be seen as a detector that can highly match with abnormal codebooks on a high probability.

All of detected abnormal codebooks and evolved ones are saved as detectors. The criterion of evolution can include but not limit to close Euclidean distance between constellations, the similar codebook structure, or exchange of constellation position. When the next codebook selection occurs, tested codebook is taken to match with detectors firstly. If the matching is successful, the codebook will be determined as an abnormal codebook. When matching fails, a detection based on legal codebook features is carried out. If the number of detectors evolved from one known codebook is meet the requirements, it is considered that a kind of unknown abnormal codebooks can be detected only by detectors. But it is still needed to regularly update detectors.

5. Conclusions

When there is a malicious codebook unknown to the receiver in the codebook pool of grant-free SCMA, it will seriously affect correct demodulation of data. Most malicious codebooks are unknown, and AIS can do unknown data detection because of its adaptability and evolutionary ability. In this paper, we do initial abnormal codebook test based on legal codebook features, that are concluded during the process of codebook design. If a codebook has abnormal features, it is an abnormal one. And tested abnormal codebooks can be used to generate detectors. That avoids low efficiency caused by random generation and direct tolerance with large self set in the typical algorithm NSA. To take advantages of existing detectors, part of new detectors can be generated through evolution of existing detectors. That improves deficiencies about waste of previous detector' information. In the subsequent detection, codebooks to be tested can be matched with detectors or tested by using normal codebook features. When existing detectors can cover most kinds of abnormal states, we can consider that we will get a responsible result only by detectors. To keep adaptability of detectors, regularly update of detector set is necessary.

Acknowledgements

This work was partially supported by Natural Science Foundation of China (Grant No.U1736212, No.61572334, No.61872255), Sichuan Province Key Research and Development Project of China (Grant No.2018GZ0183), Sichuan University (Grant No. YJ201933); the China Postdoctoral Science Foundation (Grant No. 2019TQ0217) the Fundamental Research Funds for the central Universities, the National Key Research and Development Program of China (Grant No.2016YFB0800600).

References

[1] Saito Y, Kishiyama Y, Benjebbour A, et al. Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access[J]. 2013, 14(6):1-5.

[2] Moshavi S. Multi-user detection for DS-CDMA communications[J]. IEEE Communications Magazine, 1996, 34(10):124-136.

[3] Ping L, Liu L, Wu K, et al. Interleave division multiple-access[J]. IEEE Transactions on Wireless Communications, 2006, 5(4):938-947.

[4] Huawei. Innovate Asia: First 5G algorithm competition [EB]. 2015.

[5] Alam, Mehmood, and Q. Zhang. "Performance Study of SCMA Codebook Design." Wireless Communications and Networking Conference IEEE, 2017.

[6] Au, Kelvin, et al. "Uplink contention based SCMA for 5G radio access." GLOBECOM Workshops IEEE, 2015:900-905.

[7] Forrest S, Beauchemin C. Computer immunology[J]. Immunological Reviews, 2007, 216(1):176-197.

[8] S. Forrest, A. S. Perelson, Allen L, et al. Self-nonself discrimination in a computer[C]. In: Proceedings of the 1994 IEEE Symposium on Security and Privacy, IEEE Computer Society, 1994.

[9] Huawei. Innovate Asia: First 5G algorithm competition [EB]. 2015.

[10] Bayesteh, Alireza, et al. “Blind detection of SCMA for uplink grant-free multiple-access.” International Symposium on Wireless Communications Systems IEEE, 2014:853-857.

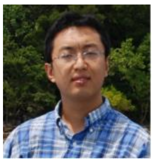
[11] Hanyuan Huang, Hui Zhao, Feilong Wang, and Jing Li, Uplink Grant-free Multi-codebook SCMA Based on High-overload Codebook Grouping, IEEE /CIC International Conference on Communications in China Workshops, 2018.

AUTHORS

Hanyuan Huang received the M.S. degree in information and Communication Engineering from Beijing University of Post and Telecommunication, Beijing, China, in 2019. She is currently pursuing the Ph.D. degree with the College of Computer Science in Sichuan University, Chengdu, China. His research interests include network security, artificial immune systems, intrusion detection, and communication engineering.

Li Tao received the Ph.D. degree in computer science from the University of Electronic Science and Technology of China, in 1994. He is currently a Professor with the College of Cyber Security, Sichuan University, China. His main research interests include network security, artificial immune systems, cloud computing, and cloud storage.

Hui Zhao received the Ph.D. degree in computer science from the Sichuan University, Chengdu, China, in 2011. He is currently an associate professor with the College of Cyber Security, Sichuan University, China. His main research interests include network security, information security, embedded system.



PARKING ASSISTANCE DISPLAY WITH ESTIMATED PARKING SPACE USING STEREO VISION

Chi-Cheng Cheng, Chi-Cheng Lee and Jyun-Han Huang

Department of Mechanical and Electro-Mechanical Engineering,
National Sun Yat-Sen University, Kaohsiung, Taiwan, R.O.C.

ABSTRACT

Inexperienced drivers always suffer from limited spatial information coming from side and review mirrors to complete parking tasks. The major obstacle is that they cannot easily estimate relative position of the parking space with respect to their vehicles. Therefore, this paper aims to develop a parking assistance display system that can continuously provide the top view of both the vehicle and the parking space for drivers. The system applies two wide-angle cameras mounted at the rear of the vehicle. In order to search for two farther corners of the parking space with efficiency, the FAST corner detection technique is employed. Three dimensional spatial coordinates of those corners can therefore be determined by the stereo vision framework. As a result, the position of the parking space relative to the vehicle can be estimated. To verify the effectiveness of the proposed parking assistance display, actual parking experiments with a golf cart were conducted. Experimental results demonstrate the parking tasks can be successfully accomplished with the help from the presented assistance display.

KEYWORDS

Parking Assistance Display, Stereo Vision, Corner Detection, Parking Space Estimation

1. INTRODUCTION

Parking is always a great challenge for drivers and strongly requires not only matured driving skills, but also accumulation of parking experience. Traditional parking assistance device consists of only the review and side mirrors. Based on the images from those mirrors, human drivers can approximately estimate environment including the parking space around the vehicle. Because of highly development of microelectronics, most parking information for drivers now comes from rear parking sensors, which give warning signal by detecting rear distance and is quite useful for parking tasks including parallel parking and garage parking. Recently, the active parking assist relying on supersonic sensor technology becomes more and more popular and the parking task can almost be accomplished without the help from drivers [1]. However, unexpected road conditions may occur anytime, human drivers still need to play the role of a supervisor especially when manual intervention on the automatic system is necessary.

The difficulty of parking stems from a number of inevitable facts. The major constraint is that the driver sits inside the vehicle's body and the clues about the environment can only be achieved by putting together pieces of image information gathered from review and side mirrors as well as visual impression from looking out of the window. Although each visual image provides partial three-dimensional environmental information, it is difficult for drivers to realize what the actual environment is based on those discontinuous and incomplete visual data.

David C. Wyld et al. (Eds): SPTM, IPPR, AIS, CSIT, DaMi, AMLA - 2020
pp. 35-44, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100604

Due to rapid progress on imaging technology, electronics fabrication, and computational capability, machine vision has been widely applied to many manufacturing areas including the automotive industry. Currently, the category of vision-based studies on automobiles probably is one of the popular research fields regarding applications of machine vision. For the purpose of preventing difficulty from parking in dark indoor site, a parking space recognition method with the light stripe projection approach was presented in [2]. To automate the target position selection of an automatic parking assist system, a parking-slot-markings recognition algorithm using two seed-points, end-points of two line segments separating the target parking slot from adjacent ones, was proposed to reduce the search range and memory requirement [3].Incorporating an around view monitor (AVM) system with a hierarchical tree structure approach, various parking slot markings can be almost fully automatic recognized [4]. Continuing from previous research, in [5], a vacant parking slot detection and tracking system using the sensor fusion technology combing an AVM system and an ultrasonic sensor was developed in 2014.

In order to offer extensive visual information about the surroundings of the vehicle, a composite top-down view of 360° of the vehicle was synthesized in [6] by combining images taken from four to six wide-angle cameras mounted around the vehicle. Similarly, an automatic parking framework based on a bird's eye vision system using four on-board fisheye cameras around the vehicle with the double circular trajectory planning and a preview control strategy was also presented [7]. In [8], a portable real-time informational bird's eye view system for the advanced driver assistance system was actually implemented. Furthermore, a parking assistance system with the depth map of the observed environment computed from dense motion-stereo was proposed to allow drivers to be able to visualize the environment around the host vehicle [9]. Compared with the ultrasound approach, the proposed approach demonstrated better accuracy and reliability in terms of both quantitative and qualitative results. Although vision-based approaches for automatic parking tasks in the field of intelligent vehicles gain interests from both academic institutions and automotive industry, there was no public benchmark dataset available for evaluation of parking-slot detection algorithms. Fortunately, a large-scale parking-slot image database comprising 8600 surround-view images collected from indoor and outdoor parking sites was established [10]. Besides, at the same time, a learning-based parking-slot detection approach was also proposed.

If estimated position of the park space relative to the vehicle can be displayed for the drivers using extra cameras, it would be helpful for drivers to conduct parking tasks since human drivers are aware of where the parking space is located. The stereo vision system is able to offer the depth information of objects in three-dimensional scene and has the cost effective advantage than other vision devices such as 360-degree panorama cameras. Besides, although the 360-degree panorama cameras are capable of delivering whole visual information around the vehicle, images for objects in the environment always suffer from severe distortion and poor image quality. Consequently, a binocular vision system installed on the back of the vehicle will be applied for position estimation of the parking space with respect to the vehicle in this study.

2. CORNER DETECTION FOR PARKING SPACE

In order to provide the position information for the parking space, an effective image processing algorithm for locating the parking space plays a crucial role. The algorithm needs to be computational efficiency to meet the requirement of fast estimation for parking tasks.

The parking space for regular cars has a shape of rectangular with the length of 5~5.7 meters and the width of 2~2.5 meters. The best feature for parking space detection should be its four corners. The FAST (Features from Accelerated Segment Test) method developed by Rosten and

Drummond in 2006 provides a time efficient algorithm for corner detection [11]. This approach starts with examining intensity difference between 16 adjacent pixels around a mask e.g., 7x7 over the target point. If the number of pixels that fulfil a given threshold value is greater than a certain number, e.g., 9, then the target point should be considered as a point at the corner. The OpenCV (Open Source Computer Vision Library), an open source computer vision and machine learning software library, offers a number of different masks for high-speed corner detection according to the FAST concept.

Since the stereo cameras are installed on the back of the vehicle, estimation of the parking space will be achieved by examining those two corners of the parking space far away from the vehicle due to the fact that they are easily captured by the cameras. The TYPE_7_12 FAST function, a 5x5 mask with 7 out of 12 pixels as depicted in Figure 1, was chosen for corner detection in small area. Furthermore, detection sensitivity is improved by applying 5 out of 12 pixels instead of regular 7 out of 12 pixels to prevent from unsuccessful detection for the left corner being an obtuse angle. After the FAST corner detection algorithm, Figure 2 demonstrates pixels for corner candidates.

In order to select appropriate inner corner point among many corner candidates from the FAST corner detection algorithm, three criteria are established. Take the right corner as an example shown in Figure 3. The first criterion is that if the point P is an inner corner point, the following conditions have to be satisfied.

$$G_3 > G_P \text{ and } G_3 > G_9 ,$$

Where G_i represents the gray scale of the i-th pixel. The second criterion is that the points with significantly different angular style need to be removed. Figure 4 illustrates a right corner described by pixels 69 and 89 in a 112-pixel circular mask. With ±4 pixels tolerance, only points within the ranges of 65-73 and 85-93 will be reserved for further inner corner verification. After passing the first two criteria, the remaining pixels are always located around the inner corner. The unique inner corner pixel can be easily obtained with the final criterion of the smallest included angle.

While the driver is performing the parking task, the scene grabbed by the cameras is always different from that taken at previous time instant. In order to enhance computational efficiency for corner detection, the detected inner corner from previous time instant will be applied as the center of ROI (Region of Interest) for searching for inner corners of the parking space at the current time instant. Besides, the angular style solved from criteria 1 and 2 at previous moment will be directly used as the initial condition for the current instant.

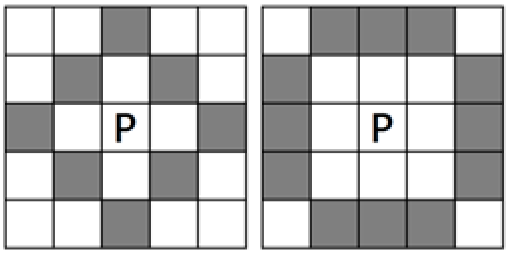


Figure 1. TYPE_5_8 and TYPE_7_12 masks in OpenCV for FAST corner detection

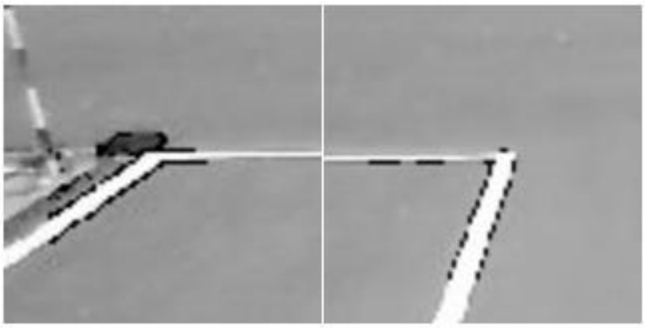


Figure 2. Corner candidates after FAST corner detection algorithm

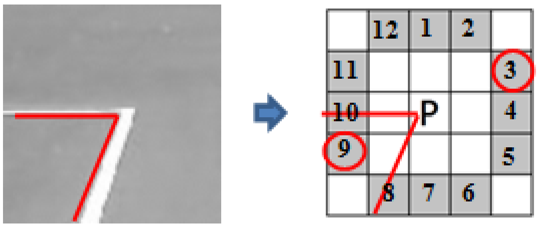


Figure 3. The first criterion for inner corner detection

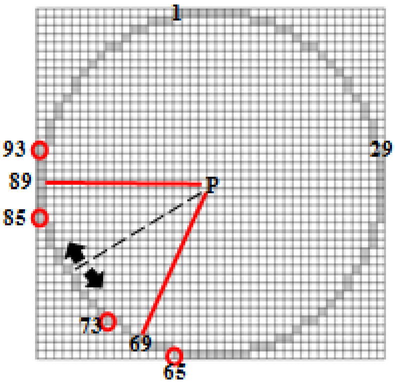


Figure 4.The second criterion for inner corner detection

3. Estimation of the Parking Space

After successful corner detection described in previous section, the three-dimensional coordinates of the parking space will be determined by a binocular vision framework. The binocular vision framework provides two images at the same time and is able to solve depth information of the target point using the clue of disparity defined as the difference between project points on two image planes.

Let two identical cameras be placed laterally with a baseline distance D and their optical axes are parallel with each other as illustrated in Figure 5. Based on the model of the stereo vision, the projection points (u_R, v_R) and (u_L, v_L) on the right and left image planes respectively can be expressed by

$$u_R = \alpha_R \frac{x_R}{z_R} + u_{0R}\,,\, v_R = \beta_R \frac{y_R}{z_R} + v_{0R}\,,$$

$$u_L = \alpha_L \frac{x_L}{z_L} + u_{0L}\,,\ v_L = \beta_L \frac{y_L}{z_L} + v_{0L}\,,$$

Where α_i and β_i are products of the focal length and the aspect ratio in horizontal and vertical directions due to non-square pixels, respectively. Besides, (u_{0R}, v_{0R}) and (u_{0L}, v_{0L}) stand for the offsets between the image center and the principal point because of imperfect assembly of image sensors. Those intrinsic camera parameters can be solved by standard calibration procedures.

After a number of mathematical manipulations, the target position (x_R, y_R, z_R) with respect to the right camera can be obtained as

$$x_R = \frac{\alpha_L D(u_R - u_{0R})}{-\alpha_L(u_R - u_{0R}) + \alpha_R(u_L - u_{0L})},$$

$$y_R = \frac{v_R - v_{0R}}{\beta_R} \cdot \frac{(\alpha_R - \alpha_L)x_R - \alpha_L D}{u_R - u_{0R} - (u_L - u_{0L})},$$

$$z_R = \frac{(\alpha_R - \alpha_L)x_R - \alpha_L D}{u_R - u_{0R} - (u_L - u_{0L})}$$

Similarly, the same target position but in the left camera coordinate system (x_L, y_L, z_L) can also be derived as

$$x_L = \frac{\alpha_R D(u_L - u_{0L})}{-\alpha_R(u_L - u_{0L}) + \alpha_L(u_R - u_{0R})},$$

$$y_L = \frac{v_L - v_{0L}}{\beta_L} \cdot \frac{(\alpha_L - \alpha_R)x_R + \alpha_R D}{u_L - u_{0L} - (u_R - u_{0R})},$$

$$z_L = \frac{(\alpha_L - \alpha_R)x_R - \alpha_R D}{u_L - u_{0L} - (u_R - u_{0R})}.$$

Ideally, $x_R = x_L - D$, $y_R = y_L$, and $z_R = z_L$. Due to limited image resolution and inevitable noises, those ideal equations cannot be perfectly satisfied. But the target position can still be successfully estimated. For future implementation, a new reference system with its origin at the middle point of origins of both right and left camera coordinate systems is established. Then the target point (x_C, y_C, z_C) with respect to the new reference system as illustrated in Figure 5 can be simply formulated by

$$x_C = \frac{x_R + x_L}{2},$$

$$y_C = \frac{y_R + y_L}{2},$$

$$z_C = \frac{z_R + z_L}{2}.$$

In addition, for the purpose of covering the parking space on the ground, both cameras need to be pointed downwards with an inclination angle θ with respect to the horizontal level as shown in Figure 6. In order to determine the relative location of the parking space with respect to the vehicle in the global reference frame, the position information extracted by the cameras need to be transformed to a reference system (*X*, *Y*, *Z*) using

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_C \\ y_C \\ z_C \end{bmatrix}.$$

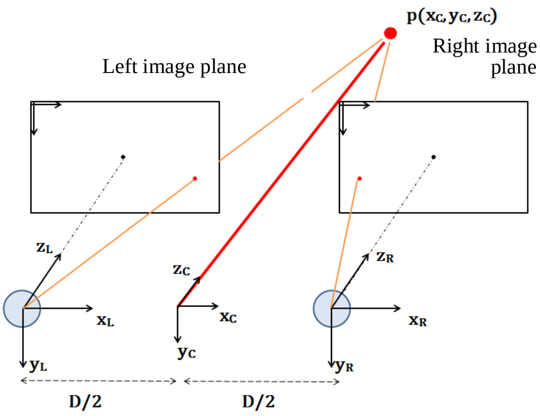


Figure 5. Reference frames for the binocular system

The origin of the reference system (*X*, *Y*, *Z*) coincides with that of the reference system (x_C, y_C, z_C), but with its *Y*-axis pointing downwards, and both *X* and *Z* axes parallel with the ground. It is possible all four corner points may not be captured by the cameras during the parking process especially when the vehicle is close to the parking space. Fortunately, since the dimension of the parking space can be already known, estimation of the parking space can still be accomplished by only two farther corner points.

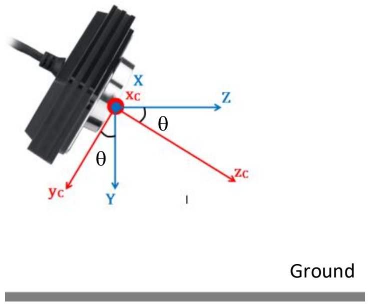


Figure 6. The camera pointing downwards with an inclination angle

4. Parking Experiments

In order to verify performance of the proposed parking assistance display, actual parking experiments with a golf cart were conducted in the campus. The golf cart shown in Figure 7 is equipped with a laptop computer, which generates the parking assistance display to drivers, and a dual-camera system on its back. The golf cart has a length of 240 cm and a width of 102 cm. The cameras are installed on top of the vehicle with a height of 160 cm and an inclination angle of 30 degrees downwards. The baseline between those two cameras is 20 cm. The camera chosen for experiments is a Widecam 1050 by Genius, which is a CMOS device that owns the advantage of low energy dissipation for portable applications. The camera has a resolution of 640x480 pixels and a wide view angle so that the parking space can be easily captured. All image processing algorithms and the assistance display were performed by a Lenovo t530 laptop computer accompanied with Visual Studio and OpenCV as software development tools. The target parking space is 475 cm long and 214 cm wide.

Actual parallel parking experiments were conducted to examine the parking performance with the help from the proposed parking assistance display. For the purpose of preventing the driver from acquiring information around the vehicle using eyes, a black curtain covering the vehicle's body was applied to block the view of the driver. The only clue for performing the parking task was the assistive display shown on the screen of the laptop as depicted in Figure 8. Figure 9 illustrates responses of the relative distance and orientation angle between the vehicle and the parking space, which are defined in Figure 10, for a parking experiment. Figure 11 presents the relationship between the estimated and the actual parking space at 5.8 second. Apparently, position error existed due to imperfect estimation of the location of the parking space. The deviation mainly came from the error caused by inaccurate depth estimation. Furthermore, during parking experiments, the driver seemed being disturbed by inconsistent images of the parking space. Although the location of the parking space could not be exactly estimated, the proposed parking assistance display still provided important information regarding relative position of the vehicle with respect to the parking space. In addition, the estimation accuracy improved when the vehicle got closed to the parking space. As a result, the driver was able to accomplish the parking task just relying on the parking assistance display proposed by this research.

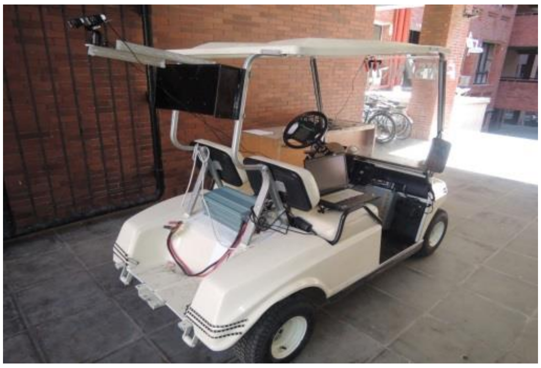


Figure 7. The golf cart with stereo vision for actual parking experiments

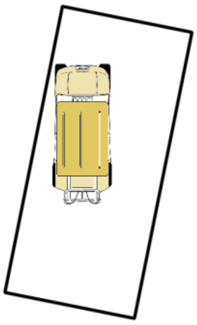


Figure 8. The parking assistance display shown on the screen

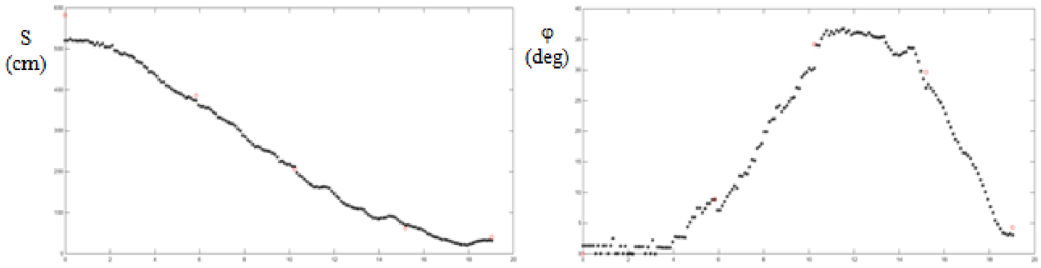


Figure 9. Distance and orientation responses of a parallel parking experiment

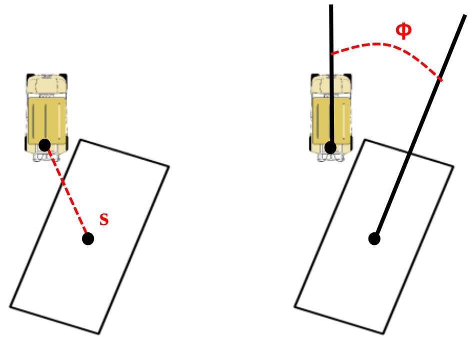


Figure 10. Illustrative distance and orientation for parking experiments

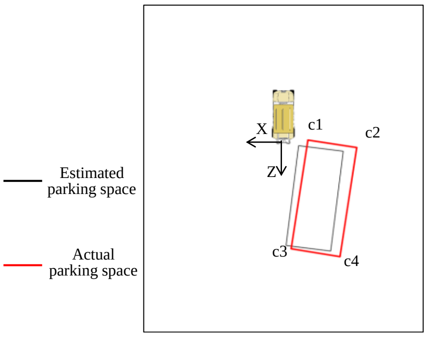


Figure 11. Estimated and actual parking space for a parking experiment at 5.8 second

5. CONCLUSIONS

The key issue for drivers to complete the parking mission is that they need to know the surroundings around the vehicle especially for the parking space as accurate as possible. For inexperienced drivers, it is very difficult to combine incomplete and fragmented images from side and review mirrors to establish spatial representation for the environment of the vehicle. Therefore, a parking assistance display based on stereo vision by two wide-angle cameras mounted at the rear of the vehicle is proposed to help drivers conducting parking tasks. Based on images from those two cameras, the relative position of the parking space with respect to the vehicle can be calculated. Furthermore, an inside-out display to show the top view of visual information about the parking space can therefore be drawn. Actual parking experiments with a golf cart were conducted to examine parking performance with the presented assistance display.

Experimental results clearly demonstrate drivers are able to accomplish parallel parking tasks only with the visual information from the parking assistance display.

For future works, a number of research directions need to be deeply explored. The proposed approach has to extend to garage parking or reverse parking tasks in the near future. Obviously, a more accurate estimation algorithm for position of the parking space is required to provide consistent estimation results. In addition, the estimation algorithm for parking space has to be robust enough to overcome the difficulty caused by its corner features occluded by other objects such as adjacent vehicles.

REFERENCES

[1] Noh, B. Tae. & Kim, S. Jae (2015) Smart Parking Assist System of Vehicle and Control Method Thereof, United States Patent 9,168,955 B2, filed March 21, 2013, and issued October 27, 2015.

[2] Jung, G. Ho, D. Kim, S. Dong & Kim, J. (2010) “Light-stripe-projection-based target position designation for intelligent parking-assist system”, IEEE Trans. Intelligent Transportation Systems, Vol. 11, No. 4, pp 942-953.

[3] Jung, G. Ho, Kim, S. Dong, Yoon, J. Pal & Kim, J. (2008) “Two-touch type parking slot marking recognition for target parking position designation”, Proc. IEEE Intelligent Vehicles Symposium, Eindhoven, The Netherlands, pp 1161-1166.

[4] Suhr, K. Jae & Jung, G. Ho (2012) “Fully-automatic recognition of various parking slot markings in around view monitor (AVM) image sequences”, Proc. IEEE Intelligent Transportation Systems, Anchorage AK, USA, pp 1294-1299.

[5] Suhr, K. Jae & Jung, G. Ho (2014) “Sensor fusion-based vacant parking slot detection and tracking”, IEEE Trans. Intelligent Transportation Systems, Vol. 15, No. 1, pp 21-36.

[6] Zhang, B., Appia, V., Pekkucuksen, I., Batur, U. Aziz, Shastry, P., Liu, S., Sivasankaran, S., Chitnis, K. & Liu, Y. (2014) “A surround view camera solution for embedded systems”, Proc. IEEE Computer Vision and Pattern Recognition Workshops (CVPRW ‘14), Columbus OH, USA, pp662-667.

[7] Wang, C., Zhang, H., Yang, M., Wang, X., Ye, L. & Guo, C. (2014) “Automatic parking based on a bird’s eye view vision system”, Advances in Mechanical Engineering, Vol. 2014, Article no. 847406, 13 pages.

[8] Gojak, V., Janjatovic, J., Vukota, N., Milosevic, M., Bjelica, Z. Milan (2017) “Informational bird's eye view system for parking assistance”, Proc. IEEE 7th International Conference on Consumer Electronics, Berlin, Germany, pp 103-104.

[9] Unger, C., Wahl, E. & Ilic, S. (2014) “Parking assistance using dense motion-stereo”, Machine Vision and Applications, Vol. 25, pp 561-581.

[10] Zhang, L., Li, X., Huang, J., Shan, Y. & Wang, D. (2018) “Vision-based parking-slot detection: A benchmark and a learning-based approach”, Symmetry, Vol. 10, No. 64, 18 pages.

[11] Rosten, E. & Drummond, T. (2006) “Machine learning for high-speed corner detection”, Proc. 9th European Conference on Computer Vision (ECCV ’06), Lecture Notes in Computer Science, Vol. 3951. Springer, Berlin, Heidelberg, pp 430-443.

AUTHORS

Chi-Cheng Cheng was born in Taipei, Taiwan, R.O.C. He received the B.S. and M.S. degrees in power mechanical engineering from National Tsing Hua University, Hsinchu, Taiwan, in1981 and 1983, respectively, and the Sc.D. degree in mechanical engineering from Massachusetts Institute of Technology, Cambridge, MA, USA in 1991. He has been with National Sun Yat-Sen University in Taiwan since 1991 and is a Professor in Department of Mechanical and Electro-Mechanical Engineering. He was a Visiting Scholar in the Department of Electrical and Computer Engineering of University of British Columbia, Canada in 2002 and a Visiting Professor with the School of Engineering Science in Simon Fraser University, Canada in 2009. He was the Dean of Office of International Affairs from 2007 to 2009, the Chairman of the Department from 2014 to 2017, and Currently the Vice Dean of College of Engineering of National Sun Yat-Sen University. His research interests are in system dynamics and control, Machine vision, mechatronics, intelligent robots, and man-machine interface. Dr. Cheng has published more than 150 technical articles in refereed international journals, conferences and book chapters. He has won the Excellent Research Award from National Science Council of Taiwan in 1998, 2001, and 2002. He has been a member of Institute of Electrical and Electronics Engineers (IEEE) since 1987. He has also been a senior member of International Association of Computer Science and Information Technology (IACSIT) since 2011.

MELANOMA DETECTION IN HISTOPATHOLOGICAL IMAGES USING DEEP LEARNING

Salah Alheejawi, Richard Berendt, Naresh Jha and Mrinal Mandal

University of Alberta, Edmonton, Alberta, Canada

ABSTRACT

Histopathological images are widely used to diagnose diseases such as skin cancer. As digital histopathological images are typically of very large size, in the order of several billion pixels, automated identification of abnormal cell nuclei would be very helpful for doctors to perform fast diagnosis. In this paper, we propose a technique, using deep learning algorithms, to first segment the cell nuclei in Hematoxylin and Eosin (H&E) stained images and detect the abnormal melanocytes on the histopathological images. The cell segmentation is done by using a novel Convolutional Neural Network (CNN) architecture. The segmented cells are then classified into melanoma and other nuclei using a Support Vector Machine classifier. Experimental results show that the CNN can segment the nuclei with more than 90% accuracy. The proposed technique has a low computational complexity.

KEYWORDS

Histopathological image analysis, Nuclei segmentation, Melanoma Detection, Deep learning.

1. INTRODUCTION

Melanoma is an abnormal growth of melanocytes which mainly occurs on the skin and it can transfer to any part of the body and destroys the tissue. About 7,800 Canadians have been diagnosed with melanoma skin cancer in 2019 and 1,300 of them would be in a fatal stage [1]. The early diagnosis of melanoma is very important as it helps to increase the chances of successful treatment and the survival rate. The Computer-aided diagnosis (CAD) techniques can effectively help doctors to diagnose and detect the melanoma in early stages [2]. The digitized histopathological slides, which are typically obtained by staining and scanning the biopsy slides of the skin tissue, can provide the cell morphological features with a high resolution. The digitized slides are known as Whole Slide Images (WSIs) and with help of CAD techniques that will permit the pathologist for precise diagnosis [3]. Pathologists usually use H&E stained images, because the morphological features of the melanocytes and other cells become vividly clear. In H&E stained image, the cell nuclei contain chromatin and that can be observed in blue shade while the cytoplasm and other connective tissues are observed with varying shades of pink. Fig. 1 shows the section of skin tissue divided into three layers: epidermis, dermis and subcutaneous layer. Due to the large density of cell nuclei in the epidermis layer, the tissue appears in dark purple color. Fig. 1 shows skin tissue image and zoomed patch obtained with H&E stain, where the abnormal melanocytes appear with irregularity in shape and color intensity [4,5].

David C. Wyld et al. (Eds): SPTM, IPPR, AIS, CSIT, DaMi, AMLA - 2020
pp. 45-53, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100605

Several techniques have been proposed to segment the cell nuclei in histopathological images [6-11]. Xu et al. [6] proposed an automated technique (henceforth referred to as the Watershed+Voting technique) to segment the cell nuclei in H&E stained images. The technique detects the nuclei seeds by using voting areas and segments the nuclei cells using marked watershed algorithm. The technique provides a good performance with high computational complexity due the seed detection algorithm. Xu et al. [11] also proposed cell nuclei segmentation technique (henceforth referred to as the gLoG+mRLS technique) using generalized Laplacian of Gaussian (gLoG) filters to detect the seeds nuclei and multiple Radial Lines Scanning (mRLS) algorithm to segment the cells. The mRLS uses high gradient pixel locations and shape information to accurately segment the cell nuclei.

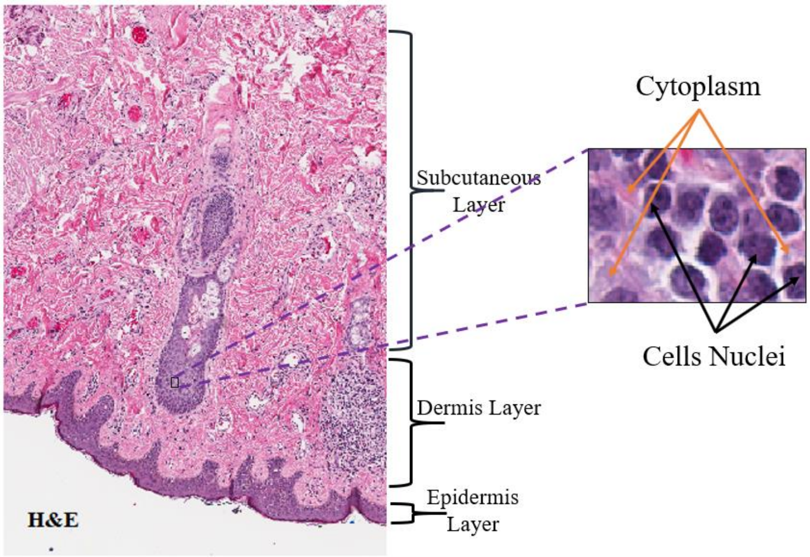


Figure 1. Example of digitized biopsy images of skin tissue image with H&E stain.

The techniques [6-11] mentioned above are generally based on extracted hand-crafted features that require significant time to process. The deep learning algorithms using CNN have been widely used successfully in medical image analysis. The CNN models can train the feature extraction process to provide high performance with low computational complexity in many different tasks (e.g. classification, detection or segmentation [12-13]). Basrinarayanan et al. [14] proposed the SegNet architecture for object segmentation. The architecture uses a number of sampling and upsampling layers for extracting the features in hierarchical levels. Ronneberger et al. [15] proposed the U-Net architecture for biomedical image segmentation. The U-Net architecture has encoder and decoder sides with number of sampling and upsampling layers, respectively. The upsampling layer outputs are enhanced by concatenating them with features from the encoder side.

In this paper, we proposed an automated technique to segment the cell nuclei and classify them into melanoma and other nuclei using a Support Vector Machine (SVM) classifier. The technique uses a CNN architecture to segment the cell nuclei on H&E stained images. The proposed CNN used several convolutional layers with different size of filters. Experimental results demonstrate high accuracy and low computational complexity of the proposed technique compared to the state-of-the-art techniques.

The organization of the paper is as follows. Section 2 describes the dataset used to train and evaluate the proposed technique. Section 3 describes the proposed technique in detail. Section 4 presents the performance evaluation, followed by the conclusion in Section 5.

2. Data Description

In this section, the nuclei segmentation and cell classification will be evaluated based on digitized biopsy slides generated at the University of Alberta. The biopsies were collected at the Cross Cancer Institute, Edmonton, Canada in accordance with the protocol for the examination of specimens with skin melanoma. The formalin-fixed paraffin-embedded tissue blocks of these biopsies were cut into thin slices (e.g., 4μm for light microscope). These slices were then mounted to glass slides and stained using H&E stain [5]. The WSIs were obtained by scanning the H&E slides using aperio scanscope scanning system under 40X magnification. The size of a WSI is typically around 40,000×60,000 pixels (in color) and each WSI contains thousands of cell nuclei. The image dataset consists of 64 WSIs for skin tissue and 9 WSIs for lymph node tissue.

3. Proposed Technique

The schematic of proposed technique is shown in Fig. 2 which consists of two modules: CNN-based nuclei segmentation and nuclei classification. The details of each module are presented in the following.

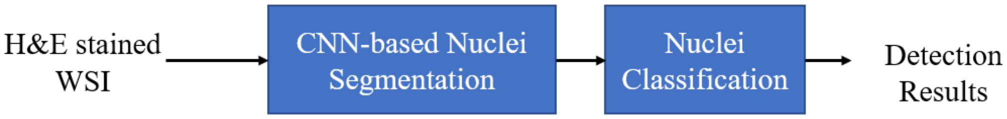


Figure 2. Schematic of a melanoma detection technique.

3.1. CNN-Based Nuclei Segmentation

In this module, the input H&E stained images are segmented into cell nuclei and background. The nuclei segmentation is done by using the proposed CNN architecture, henceforth referred to as the NS-Net architecture (Nuclei Segmentation Net). The NS-Net architecture, shown in Fig. 3, consists of five convolutional layers (shown in gray color) and one softmax (shown in pink) followed by the pixel classification layer (shown in blue). The convolutional layer in the NS-architecture consists of three operations: convolution, batch normalization [16], and Rectified Linear Unit (ReLU) activation [17]. A brief description of each operation is presented in the following:

(i) Convolution: In this layer, the input image *I* is convolved with filters F_j :

$$R_j = I * F_j, \quad j = 1, 2, .., N \tag{1}$$

where *N* is the number of filters, R_j is the output corresponding to the *j*th convolution filter, and F_j is the weights of the *j*th filter.

(ii) Batch normalization: This operation is used to normalize the convolutional layers output to zero mean and unitary variance across a current mini-batch. The normalized $\hat{R}_i$ output will be scaled with σ and shifted by β as follows:

$$y_i = \sigma \hat{R}_i + \beta \qquad (2)$$

(iii) ReLU: It is an activation function that will output the input values *y* that are greater than zero using the following equation

:

$$f = \max(0, y) \qquad (3)$$

where *f* is output of the activation function ReLU.

In this architecture, the features are extracted in hierarchical levels by using convolutional filters of different sizes. The change on the convolutional filters can precisely locate the object boundaries that need to be segmented. Most existing CNN architectures include pooling layers. In our experiment, it has been found that the pooling leads to loss of the spatial information that carries important texture and shape features of the nuclei. Therefore, the pooling layer has been omitted in the proposed architecture. Table 1 shows the number and the size of filters in each layer of the NS-Net architecture. The NS-Net architecture is trained and evaluated using a dataset of 24 high resolution H&E stained images (1920×2500 color pixels). Each image is divided into overlapping blocks of 64×64 color pixels to obtain 458 block-images. The block-image dataset is divided into 70% for training, 15% for testing and 15% for validation. The entropy loss function with the stochastic gradient descent with momentum (SGDM) optimizer is used to train the NS-Net architecture [18]. Fig. 4 (a) shows an input H&E stained image and (b) shows the masked nuclei image obtained using the NS-Net architecture.

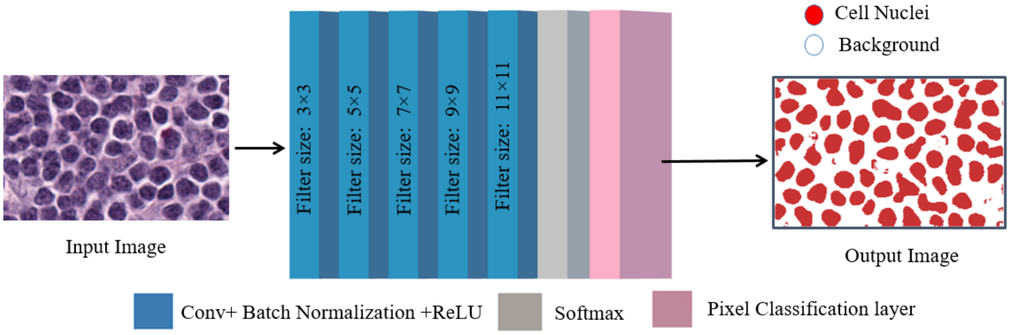


Figure 3. The NS-Net architecture for nuclei segmentation (with 5 convolutional layers).

Table 1. Details of the NS-Net architecture with 5 convolutional layers.
Input image size: M×N pixels (color). Number of classes: C.

	Number of Filters	Number of Channels	Output Image size	Filter size
Layer-1	64	3	M×N×64	3×3
Layer-2	64	64	M×N×64	5×5
Layer-3	64	64	M×N×64	7×7
Layer-4	64	64	M×N×64	9×9
Layer-5	C	64	M×N× C	11×11
Softmax layer	-	C	M×N×C	

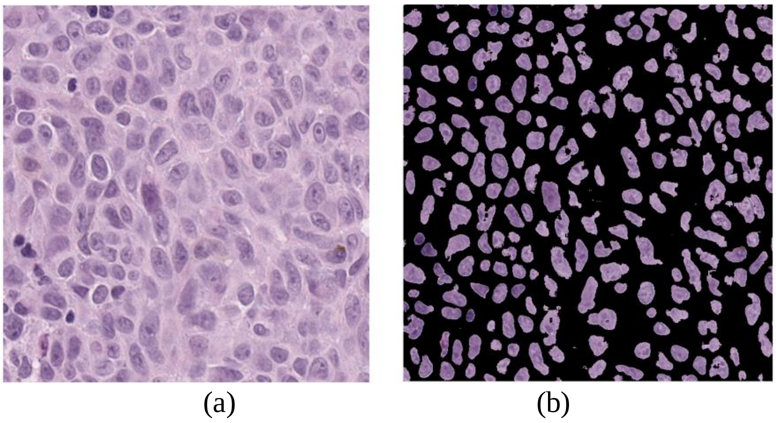


Figure 4. Segmentation results. (a) Input image (b) Segmented image obtained using the NS-Net architecture.

3.2. Nuclei Classification

In this module, the segmented nuclei obtained using the NS-Net architecture is classified into two classes based on hand-crafted features. The feature vector consists of 18 first-order features, 9 Histogram of Oriented Gradient features, 24 Haralick texture features and 3 Morphological features. The features are extracted for each pre-segmented nuclei and described briefly as follows:

(i) First-order features: It includes histogram-based features such as the mean, standard deviation, third moment, smoothness, entropy, and uniformity for 3-channels (R, G and B) to obtain 18 features (6×3).

(ii) Histogram of Oriented Gradient features: It measures the gradient of (9) orientations in localized portions of the segmented nuclei image.

(iii) Haralick texture features: It is calculated from a Gray Level Co-occurrence Matrix, (GLCM). It includes six GLCM features (correlation, energy, homogeneity, contrast, entropy, and dissimilarity) in 4 directions (i.e., 0°, 45°, 90° and 135°) to obtain 24 features (6×4).

(iv) Morphological features: It includes the eccentricity, solidity, and the ratio of major and minor axes of the cell nuclei to obtain (3) features.

The extracted feature vector (size 54) of each cell nuclei are then classified into normal and melanoma using SVM classifier [19-20]. The SVM is a very efficient supervised classifier that can handle even a non-linearly separable features and create hyperplane to separate melanocyte from other cells. In the proposed technique, the SVM model is trained and tested on 800 cell nuclei (70% for training and 30% for testing). In this technique, the SVM classifier is applied with different kernels and it shown that the Gaussian kernel provides the best classification results (see Table 4 in section 4).

After the nuclei in the lymph node are classified in the lymph node, the staging of melanoma is done based on the number and size of melanocytic clusters present in the lymph node image.

4. RESULTS AND DISCUSSIONS

In this section, we present the performance of the proposed technique using the dataset mentioned in Section 2.

The segmentation performance is evaluated and compared with handcrafted feature-based algorithms such as gLoG+mRLS and Voting+Watershed techniques and with trained CNN features techniques such as SegNet and U-Net architecture. The segmentation performance is evaluated using Accuracy, Precision, Recall and BF-score [21] measures defined as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \times 100\%$$

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

$$\mathrm{Recall} = \frac{TP}{TP + FN} \times 100\%$$

$$BF\ Score = \frac{2TP * TN}{TP * TN + 2TN * FP + TN * FN} \times 100\%$$

where *TP*, *TN*, *FN* and *FP* denote the number of true positives, true negatives, false negatives and false positives, respectively. Table 2 shows the segmentation performance of different techniques. It is observed that the deep learning algorithms provide excellent performance compared to the classical feature-based algorithms. This is because the classical features are less sensitive to the diversity of the cell nuclei in the skin tissue. For example, the abnormal melanocytes tend to have light and inhomogeneous color (see Fig. 5) and that causes miss detection of the melanocytes in the gLoG+mRLS and Voting+Watershed techniques.

Table 2. Segmentation Performance of the deep learning algorithms and the classical feature-based algorithms.

Technique:	Accuracy	Precision	Recall	BF-Score	Execution time (in s)
Voting+Watershed [6]	83.64	78.24	84.64	81.31	143.71
gLoG+mRLS [11]	76.67	79.27	60.25	68.46	128.57
SegNet [14]	87.84	84.16	87.53	85.81	15.37
U-Net [15]	78.79	87.41	57.87	69.63	20.82
NS-Net	90.21	87.20	89.90	88.52	14.27

In this work, the NS-Net, SegNet and U-Net architectures are trained with the same number of training images. The NS-Net architecture is also evaluated with CNN architecture in terms of the required parameters need to be train as shown in Table 3.

Figs. 6 (b)-(f) shows the subjective segmentation performance of SegNet [14], U-Net [15], gLoG+mRLS [11], Voting+Watershed [6] and the NS-Net architecture, respectively. It is observed that the NS-Net architecture provides excellent nuclei segmentation, whereas gLoG+mRLS, Voting+Watershed techniques miss a few cell nuclei due to the inhomogeneity in the cell nuclei color. It is also observed that the U-Net architecture does not perform well compared to the other techniques because the overfitting due the large number of the filters that are used in the cell nuclei segmentation.

Table 3. Properties of CNN architectures used in performance evaluation.

CNN Architecture	Convolutional layers	No. of Trained parameters	Filter size	No. of Filters
SegNet [14]	8	225,542	3×3	64
U-Net [15]	11	905,472	3×3	(64, 128, 256)
NS-Net	5	150,336	(3×3)- (11×11)	64

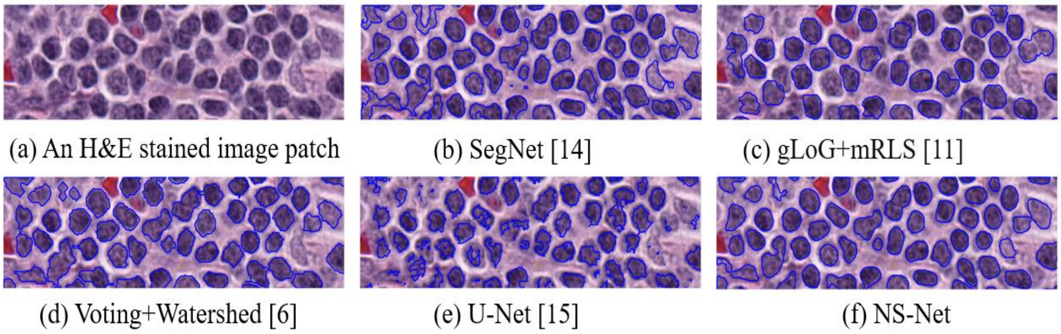

(a) An H&E stained image patch
(b) SegNet [14]
(c) gLoG+mRLS [11]
(d) Voting+Watershed [6]
(e) U-Net [15]
(f) NS-Net

Figure 5. Subjective comparison of cell nuclei segmentation results (contoured in blue color) (a) original test image, (b)-(f) Segmentation results for SegNet [14], gLoG+mRLS [11], Voting+Watershed [6], U-Net [15] and NS-Net techniques, respectively.

The classification performance is evaluated in terms of the Accuracy, Precision, Recall and BF Score measures. The SVM classifier has been evaluated with different kernels (such as Gaussian, linear and polynomial kernels [18]). The results are shown in Table 4. It is observed that the SVM classifier with Gaussian kernel provides the best performance for the dataset. Fig. 6(a)-(b) shows the nuclei classification results obtained using Gaussian kernel, where the melanoma and other nuclei are contoured in red and blue colors, respectively.

Table 4. Performance of the nuclei classification using different SVM kernels.

Evaluation Measures	SVM Kernel		
	Linear	Polynomial	Gaussian
Accuracy	80.52	57.28	85.72
Precision	77.43	57.28	80.04
Recall	93.14	96.65	97.42
BF Score	84.56	71.93	87.87

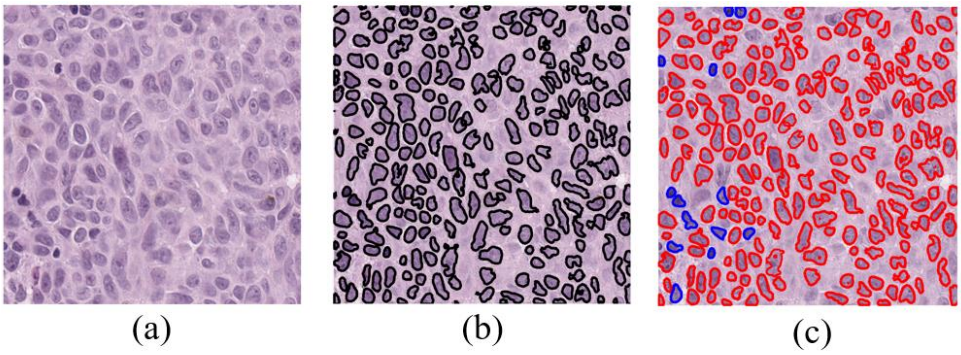


Figure 6. Example of classification results. (a) NS-Net input image (b) NS-Net segmented output image (c) Classified image obtained using SVM, where melanoma and other cell nuclei are contoured with red and blue color, respectively.

5. CONCLUSION

This paper proposes an automated technique for melanoma detection on the skin/ lymph node histopathological images. The technique segments the cell nuclei in H&E stained lymph node image using the NS-Net architecture. The NS-Net architecture segments the image into background and cell nuclei regions. The proposed CNN architecture provides an excellent segmentation performance with a low computational complexity. The segmented cell nuclei are then classified into melanocytic and other cell nuclei using an SVM classifier with Gaussian kernel. After the nuclei classification, the staging of melanoma can be done by the doctors based on the number and size of melanocytic clusters present in the image.

REFERENCES

[1]. Rebecca L Siegel, Kimberly D Miller, & Ahmedin Jemal (2019), "Cancer Statistics, 2019," CA: A Cancer Journal for Clinicians, Vol. 69, No. 1, pp.7–34.

[2]. Lieve Brochez, Evelien Verhaeghe, Edouard Grosshans, Eckhart Haneke, Gérald Piérard, Dirk Ruiter, & Jean-Marie Naeyaert (2002), "Inter-observer variation in the histopathological diagnosis of clinically suspicious pigmented skin lesions". J. Pathol. 196(4), pp.459–466.

[3]. Cheng Lu, Muhammad Mahmood, Naresh Jha, & Mrinal Mandal, (2013), "Automated segmentation of the melanocytes in skin histopathological images". IEEE J. Biomed. Health Inf. 17(2), pp. 284–296.

[4]. Marcial García Rojo, Ana Morillo Castro, & Luis Gonçalves, (2011), "COST action "eurotelepath": digital pathology integration in electronic health record, including primary care centres," Diagnostic Pathology 6, S6 (2011), pp. S1–6.

[5]. Ronald S Weinstein, Anna R Graham, Lynne C Richter, Gail P Barker, Elizabeth A Krupinski, Ana Maria Lopez, Kristine A Erps, Achyut K Bhattacharyya, Yukako Yagi, & John R Gilbertson, (2009) "Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for thefuture," Human pathology, vol. 40, no. 8, pp. 1057–1069.

[6]. Hongming Xu, Cheng Lu, & Mrinal Mandal (2013), "An efficient technique for nuclei segmentation based on ellipse descriptor analysis and improved seed detection algorithm". IEEE J. Biomed. Health Inf. 18(5), pp. 1729–1741.

[7]. Olcay Sertel, (2012), "Image analysis for computer-aided histopathology," Ph.D. dissertation, The Ohio State University.

[8]. Juliana M Haggerty, Xiao N Wang, Anne Dickinson, Chris J O'Malley & Elaine B Martin, (2014), "Segmentation of epidermal tissue with histopathological damage in images of haematoxylin and eosin stained human skin". BMC Med. Imaging. 14(1), pp. 7.

[9]. Jierong Cheng, Jagath C. Rajapakse, (2009), "Segmentation of clustered nuclei with shape markers and marking function," IEEE Transactions on Biomedical Engineering, vol. 56, no. 3, pp. 741–748.

[10]. Bahram Parvin, Qing Yang, Ju Han, Hang Chang, Bjorn Rydberg, Mary Helen & Barcellos-Hoff, (2007), "Iterative voting for inference of structural saliency and characterization of subcellular events," IEEE Transactions on Image Processing, vol. 16, no. 3, pp. 615–623.

[11]. Hongming Xu, Cheng Lu, Richard Berendt, Naresh Jha & Mrinal Mandal, (2017), "Automatic nuclei detection based on generalized Laplacian of Gaussian filters," IEEE Journal of Biomedical and Health Informatics (JBHI), vol.21, no.3, pp.826-837.

[12]. Jonathan Long, Evan Shelhamer, & Trevor Darrell, (2015), "Fully convolutional networks for semantic segmentation," in CVPR, pp. 3431–3440.

[13]. David Eigen, & Rob Fergus, (2015) "Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture", in Proceedings of the International Conference on Computer Vision, pp. 2650–2658.

[14]. Vijay Badrinarayanan, Alex Kendall, & Roberto Cipolla (2015). "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation." *arXiv*. Preprint arXiv: 1511.0051.

[15]. Olaf Ronneberger, Philipp Fischer, & Thomas Brox (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation." Medical Image Computing and Computer-Assisted Intervention (MICCAI). Vol. 9351, pp. 234–241.

[16]. Sergey Ioffe, & Christian Szegedy, (2015), "Batch normalization: accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167.

[17]. Vinod Nair, & Geoffrey E. Hinton (2010), "Rectified linear units improve restricted Boltzmann machines," in Proceedings of the 27th international conference on machine learning, pp. 807–814.

[18]. Christian Robert, (2014) "Machine learning, a probabilistic perspective," CHANCE, vol. 27, no. 2, pp. 62–63.

[19]. Corinna Cortes & Vladimir Vapnik, (1995), "Support-vector networks". Machine Learning. 20 (3), pp. 273–297.

[20]. Asa Ben-Hur, David Horn, Hava Tova Siegelmann, & Vladimir N Vapnik, (2001) "Support vector clustering," Journal of Machine Learning Research, vol. 2, pp. 125-137.

[21]. Gabriela Csurka, Diane Larlus, & Florent Perronnin (2013), "What is a good evaluation measure for semantic segmentation?" Proc. of the British Machine Vision Conference, pp. 32.1-32.11.

THE PARALLEL HTM SPATIAL POOLER WITH ACTOR MODEL

Damir Dobric[1], Andreas Pech[2], Bogdan Ghita[1] and Thomas Wennekers[1]

[1]University of Plymouth, Faculty of Science and Engineering, United Kingdom
[2]Frankfurt University of Applied Sciences,
Dept. of Computer Science and Engineering, Germany

ABSTRACT

The Hierarchical Temporal Memory Cortical Learning Algorithm (HTM CLA) is an algorithm inspired by the biological functioning of the neo-cortex, which combines spatial pattern recognition and temporal sequence learning. It organizes neurons in layers of column-like units built from many neurons such that the units are connected into structures called regions (areas). Layers can be hierarchically organized and can further be connected into more complex networks, which would allow to implement higher cognitive capabilities like invariant representations. However, a complex topology and a potentially high number of neurons would require more computing power than a single machine even with multiple cores or a GPU could provide. This paper aims to improve the HTM CLA by enabling it to run on multiple nodes in a highly distributed system of processors; to achieve this we use the Actor Programming Model. The proposed concept also makes use of existing cloud and server less technology and it enables easy setup and operation of cortical algorithms in a distributed environment. The proposed model is based on a mathematical theory and computation model, which targets massive concurrency. Using this model drives different reasoning about concurrent execution and should enable flexible distribution of cortical computation logic across multiple physical nodes.

This work is the first one about the parallel HTM Spatial Pooler on multiple nodes with named computational model. With the increasing popularity of cloud computing and serverless architectures, this work is the first step towards proposing interconnected independent HTM CLA units in an elastic cognitive network. Thereby it can provide an alternative to deep neuronal networks, with theoretically unlimited scale in a distributed cloud environment. This paper specifically targets the redesign of a single Spatial Pooler unit.

KEYWORDS

Hierarchical Temporal Memory, Cortical Learning Algorithm, HTM CLA, Actor Programming Model, AI, Parallel, Spatial Pooler.

1. INTRODUCTION

Currently more and more popular Artificial Neural Networks employ supervised learning based on strong mathematical principles. These principals are efficient to solve specific kind of problems, but they operate in a way, which is not very well aligned with the way the brain might work.

Similarly, Recurrent Neural Networks are increasingly closing in to model of the biological functioning of parts of the brain. Unlike the brain, which typically operates in an unsupervised way, concepts like RNN and DNN apply supervised learning techniques.

David C. Wyld et al. (Eds): SPTM, IPPR, AIS, CSIT, DaMi, AMLA - 2020
pp. 55-70, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100606

Hierarchical Temporal Memory Cortical Learning Algorithm (HTM CLA) [1] is an algorithm aiming to replicate the functioning of neocortex [2]. It incorporates a spatial pooler algorithm capable of learning spatial patterns by using simple Hebbian learning rules [3]. Every time the same or a similar input recurs, synapses between the input and columns strengthen their permanence (weight) value [4]. Similarly, synapses can "forget" a learned pattern if it does not occur for a long enough time.

Such hierarchically organized structures can also be connected into networks, which provide more cognitive capabilities like invariant representation, pattern- and sequence-recognition.
The original HTM CLA was implemented in Python as a part of the NUPIC framework developed by Numenta [5]. C++ and JAVA implementations of HTM CLA are also available.[6] Because many of the modern enterprise applications are typically implemented in .NET with an increasing demand for cross-platform (Linux, Windows and MacOS) support, an implementation of this algorithm is required in the software industry to avoid inefficient and costly bridging between frameworks. As a preceding part of this work, HTM CLA was ported to C# .NET Core. [7] The current C# .NET Core version of HTM CLA aligns with JAVA implementations (which aligned with the original Python version [1]). It supports the singe-core **Spatial Pooler** and **Temporal Memory** algorithms, which are limited in some ways. Processing of information in neurons inside of HTM is sparsely encoded as in biological neuronal circuits [8]. HTM CLA, in a nutshell, uses Hebbian learning rules [9] on binary sparse arrays represented as sequence of integers (0/1). In the context of memory consumption, the current representation of binary values in the form of integers is not the most efficient. Improving this is still work in progress and this kind of optimization is not in the focus of this paper. The assumption in the present work is rather that the current algorithm, when used with a high number of neurons, is highly CPU and RAM intensive. The simple Hebbian-Rule makes internal calculations efficient in comparison to other algorithms (i.e. back-propagation). However, the complex topology of the Spatial Pooler (SP) and the Temporal Pooler (TP) in HTM CLA with a high number of neurons and synaptic connections, internal inhibition, and boosting–algorithms, requires significantly more computing power and memory than available on a single commodity machine with multiple core processors and a GPU.

Current implementations across the mentioned frameworks maintains internally several matrices and maps (key-value pairs). For example, there is a matrix of synaptic connections between input neurons and cortical columns. To create a **Spatial Pooler** instance with 128x128 sensory neurons and 1024x1024 columns, the framework will have to create this matrix with 16,384 x 1,048,576 = 17,179,869,184 elements. In a .NET framework using a 64bit architecture operating system, the maximum possible array size of integers (32 bits) is 2,147,483,591, calculated as:

$$\frac{2^{32}}{2} - 56 = 2.147.483.591 \tag{1}$$

This is a half of the maximal integer value on 64 systems subtracted by 56, which is an internal framework overhead to hold an array. This limit depends on many factors and fortunately, can be optimized. Nonetheless even with a very efficient optimization, the limitations from using a single node only will remain.

The current architecture of HTM CLA has, in this context, two limitations of interest: a limitation of the synaptic matrix size by **available memory** and long **calculation times** required for operations on synapses. Most papers related to HTM CLA indicate experimental work with 1024, 2048, 4096 (see [10]) and 16384 columns. As an example, in a case of 4096 columns and sensory input of 1024 neurons, which corresponds to an input image of 32x32 pixels, the SP algorithm will create 4,194,304 synapses when using global inhibition. The goal is therefore to design a distributed HTM CLA, which can run on multiple nodes and operate with any number of columns (i.e. >100,000). HTM CLA is redesigned for flexible horizontal scale in highly distributed systems by using an actor model. Moreover, the concrete implementation should make usage of modern container serverless technologies. Such a model should enable the flexible distribution of computation inside of a single HTM CLA unit or connecting multiple independent HTM CLA units in collective intelligence networks and provide an alternative to deep neuronal networks. This work is the first one about the parallel HTM Spatial Pooler on multiple nodes with the Actor Programming Model. With the increasing popularity of cloud computing and serverless architectures, this work is the first step towards proposing interconnected independent HTM CLA units in an elastic cognitive network and can provide an alternative to deep neuronal networks, with theoretically unlimited scale in a distributed cloud environment. This paper specifically targets the redesign of a single Spatial Pooler unit.

Section 2 in this paper describes the current state of the Actor Programming model and the Spatial Pooler. Sections 3 and 4 describe how the Spatial Pooler was improved for scale with help of Actor Model reasoning.

2. ACTOR PROGRAMMING MODEL AND CURRENT STATE OF THE SPATIAL POOLER

2.1 Actor Programming Model

The Actor Programming Model [11] is a mathematical theory [6] and computation model, which addresses some of the challenges posed by massive concurrency. In this theory, the Actor is treated as the universal primitive of concurrent computation. An Actor is a computational unit that, in response to a message it receives, can concurrently run code. Motivation for this programming model in this work is the simple reasoning about concurrent computation. Moreover, both the HTM CLA and the Actor Model are biologically inspired models.

Designing distributed systems with this model can simplify compute balancing between actors deployed on different cores and physically distributed nodes. In this paper, a node is defined as a physical machine, which hosts multiple actors.

Because this work is related to the C# .NET Core implementation of Spatial Pooler and Temporal Memory, the Actor Model implementation must support the .NET Core platform. Following a review of several actor programming model frameworks, including Orleans Virtual Actors [12], Service Fabric Reliable Actors [13], and Akka.NET [14], it became apparent that none of them is suitable for this task. While they are very powerful, such solutions do not offer custom partitioning functionality [12] [13], or they rely on some corporate-specific cluster [13]. As a result, the approach taken was to design and implement a lightweight version of the actor model framework. The most promising framework was Akka.NET [14], but it has shown insufficient results when it comes to networking under high CPU consumption. The Actor Model Framework proposed by this paper combines RPC and API style messaging to enable an easy and intuitive implementation. Message API style is used to physically decouple nodes in the cluster, which enables easier addition of new nodes while the system is running.

Immediately after start of the application, the actor local system is instantiated. This system drives the calculation and plays a role of scatter and gather. Then the reference to the remote instance of the Actor is created, which typically implements the calculus code. The reference behaves as a proxy to the actual implementation, which is running elsewhere in the cluster. Finally, the remote calculation is started by using a generalized Ask method. This method routes the request to the Actor running on some node in the cluster, requests the calculation and awaits a result. The implementation of all distributed calculations is grouped together in a class called HtmActor; calculations are triggered by appropriate messages, i.e. *CalculateOverlap*, *AdaptSynapses*, etc. More details about calculus? can be found in section 3. The ‘cluster’, which hosts the actors, is built in a generalized manner, with no knowledge of HTM CLA, and can be used for any other kind of distributed calculation. The logic inside the cluster is used to receive messages, find actors with appropriate calculus, execute calculus and send back the result of calculation. The same code with the same configuration executes on every physical node in the cluster.

2.2 Current State of the Spatial Pooler

The original implementation of the Spatial Pooler (SP) and the Temporal Memory (TM), as originally migrated from JAVA, supports in .NET Core the single threaded execution of calculus of both algorithms. To be able to support parallel execution of HTM on multicore and multimode infrastructures, the SP algorithm was initially redesigned to support multiple cores on a single machine. Note that the optimization of the TM algorithm is work in progress and it is beyond the scope of this paper. The sensory input is defined as a set of neurons by input topology; Spatial Pooler uses an internally flattened version of input vector mapped to sensory neurons. Every input is represented as a set of values (0/1), where ***N*** is the number of all sensory neurons. This number is also known as the number of features. A flattened version of the input vector $I_k$is defined as:

$$I_k = \{0,1\}^{1x\mathrm{N}} \tag{2}$$

Columns are defined as a set of grouped cells, represented as a flat column array, where ***M*** is the total number of columns:

$$C \quad = \{c1, \mathrm{c}2, .., \mathrm{cM}\}^{1x\mathrm{M}} \tag{3}$$

Most other neuronal networks typically connect every column to every input neuron. The **Spatial Pooler** connects to a subset of input neurons. This subset is defined by receptive field (**RF**) of the column. **RF**-array is defined as a subset of all column’s synapses:

$$P_k{}^{1x\mathrm{Ckq}} | \, C_{kq\epsilon} \; \{0, ..., \mathrm{N}\text{-}1\} \tag{4}$$

The original design of SP maintains a single instance of the connection matrix$\boldsymbol{\lambda}$. This matrix specifies whether the column ***Ci*** is connected to the sensory neuron ***Ij***. Indexes ***i*** and ***j*** are in the flattened versions of columns and sensory neurons respectively.

$$\lambda = \begin{pmatrix} x11 & x12 & \dots & x1\boldsymbol{N} \\ x21 & & & \\ \dots & & & \\ xM1 & xM2 & & xMN \end{pmatrix} | \ \mathrm{xij} \in \{0,1\} \tag{5}$$

Note, that the ***Ci*** column is connected to the sensory neuron ***Ij*** if the synapse permanence value is higher than the *proximal synapse activation* threshold. More details about the mathematical formalization of HTM CLA can be found in [15].

3. PROPOSED SPATIAL POOLER FOR PARALLEL EXECUTION WITH ACTOR MODEL

To be able to save memory and partition calculus of entire column space, This matrix has been was removed from original version of **SP** and semantically restructured as a graph of columns, in order to be able to save memory and partition calculus of the entire column space.
Figure 1 shows a single column inside of the column graph.

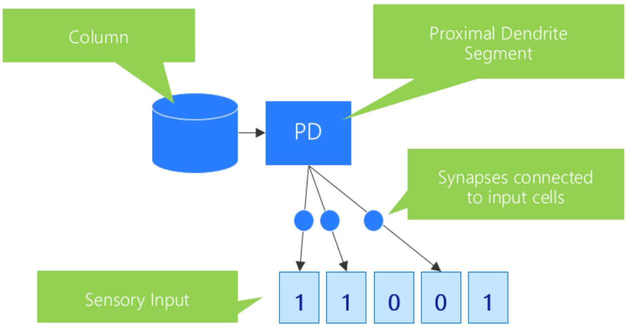


Figure 1
Representation of a single column, which can be calculated on any node in the cluster. Every column holds its own dendrite segment with a set of synapses connected to sensory input defined by its Receptive Field.

Removal of this matrix enabled easier reasoning about single column calculus, as proposed by the Actor model approach. With this, it is possible to partition columns and to share memory across multiple nodes without of need to use distributed locks, which must be used to coordinate distributed calculation. Right now, three implementations of **SP** are implemented and considered:

- *Spatial Pooler single threaded* original version without algorithm specific changes.
- *SP-MT multithreaded* version, which supports multiple cores on a single machine and
- *SP-Parallel*, which supports multi-core and multimode calculus of spatial pooler.

The Spatial Pooler algorithm consists in general of two stages inside of an application:

- Initialization
- Learning

Every named stage runs several specific calculations shown at
Figure 2. For example, the Initialization stage performs a few initialization tasks internally. The Columns and synapse initialization stage creates a graph of columns with all required artefacts. The initialization stage is typically running once, and the learning stage is running for every input sample (online-learning). SP-MT and SP-Parallel versions of SP hold the parallel implementation of all listed algorithms as shown in
Figure 2 at the right.

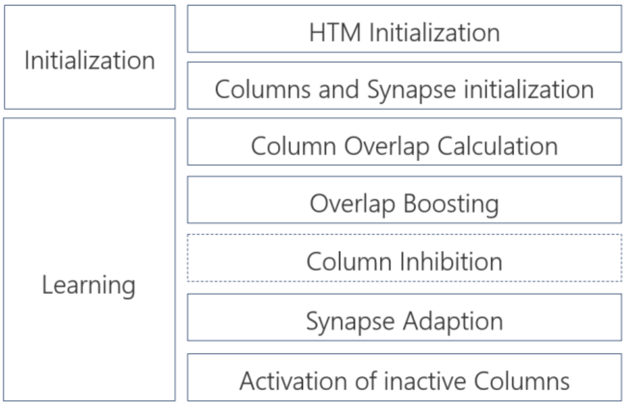

Initialization
HTM Initialization
Columns and Synapse initialization
Learning
Column Overlap Calculation
Overlap Boosting
Column Inhibition
Synapse Adaption
Activation of inactive Columns

Figure 2 : Spatial Pooler internal calculations shared across Initialization and Learning stage (left). Every stage consists of multiple calculations steps (right).

Apart from *Column Inhibition*, which is currently shared across all three implementations, all other calculations are SP-version specific. In some future versions of SP, Column *Inhibition* and some other smaller algorithmic parts (not listed here) might also be redesigned. Redesign of **SP** targets two major tasks: partitioning of memory and partitioning of CPU usage. Memory consumption inside of SP is defined as follows:

$$m \quad = m(i_k) + \textstyle\sum_u^M m_c(u) \tag{6}$$

$$m_c(u) = m_{s0} + \sum_{w}^{S} m_s(w) \tag{7}$$

where:

m – Overall memory consumption of a single SP instance, while a calculation is running.
ik – Input vector sample: $k\epsilon\{0, N - 1\}$
m(*i*k) - Memory occupied by an input sample.
m*c(u)*- Memory occupied by a single column.
ms0 - Memory occupied by column, excluding synapses. This memory is nearly the same across all columns. The difference is mainly influenced by holding references to a different number of synapses.
ms(w)- Memory inside of a column occupied by single instance of a synapse. Sum fraction of equation (7) corresponds to memory occupied by the proximal dendrite segment of a column with S synapses.

The original SP implementation and SP-MT both consume all memory ***m*** inside of a single process and it is therefore limited by physical memory of the hosting machine (node). For this reason, the first HTM Initialization step (see Figure 2) shares the same technique to allocate required memory in both named algorithms. SP-MT algorithm runs all calculations of every column on a single core by using C# technique called task/await.

4. The Partition Concept

The SP-Parallel algorithm performs partitioning of the column graph and distributes columns across multiple nodes. A partition is defined as a unit of computation with occupied memory.

$$P = p(Cpu, Mem) \tag{8}$$

Creating of partitions can be expressed by the following pseudo code:

```
createPartitions(numElements, numOfPartitions, nodes)

   destNodeIndx = 0

   numPerPart =
   round(1+numElements /
   numOfPartitions);

   FOR partIndx = 0 to numOfPartitions
   OR min>=numElements

min = numPerPart * partIndx;

   maxPartEl = numPerPart *
  (partIndx + 1) - 1;

IF maxPartEl < numElements
     max = maxPartEl
    ELSE
     max = numElements - 1;

   destNodeIndx =
   destNodeIndx % nodes.Count;

   destinationNode =
   nodes[(destNodeIndx++ %
   nodes.Count)];
  placement =
  (destinationNode,
   partIndx, min, max)

  map.Add(placement)
ENDFOR

return map;
```

This code ensures that all columns (*numOfElements*) are uniformly shared across specified *numOfPartitions* and second, that all partitions are shared uniformly across all specified nodes. For example, if numElements = 90000 (columns), nodes = 3 and number of partitions = 35 then 34 partitions will contain 2572 elements and the last partition will be filled up with the remaining 2552 elements.

To understand how SP is changed to fulfil parallelization requirements, the following SP pseudo code must be refactored:

```
compute(input, learn)
  overlap = calculateOverlap(input)
  if(learn)
     boostedOverlap = boostFactors*overlap
 else
      boostedOverlap = overlap

 activeCols = inhibitColumns(boostedOverlaps)

 adaptSynapses(input, activeCols)

 activateInactiveCols()
```

To solve this, several parallelization approaches [16] have been analysed. As a result, a dedicated and simple HTM column placement (distribution) algorithm has been designed based on the described partitioning mechanism.

Ideally, like neural parallelism [16] in the context of node parallelization, calculus for every column could be executed on a single core. For various practical reasons, placing of single column calculation on a core is understood as a special case. The partitioning algorithm rather places a set of columns in a partition, which is calculated on a single core across all nodes. In a generalized and simplified form, the overall calculation time can be defined as follows:

$$t \quad = C\mathrm{N}\, t_s + \frac{1}{N_c}\sum_{u}^{C\mathrm{N}} t_u \quad C\mathrm{N}\, t_g \mid m < m_\theta \tag{9}$$

$+$

Equation (9) states that the theoretical amount of time required to calculate any step (see Figure 2) is defined as the sum of scatter time t_s needed to remotely dispatch calculation, the sum of all column-specific calculations t_u divided by the number of cores N_c and gather time t_g needed to collect results. Note that the calculation time for every column t_u is statistically different, depending on the number of connected synapses on the dendritic segment.

This equation holds as long overall memory consumption on the node does not exceed the maximally allowed threshold ***m***θ. If this limit is exceeded, the operation system will generate hard-page faults, which would cause memory reallocation to disk. Because this operational state would dramatically slow down performance, algorithms should take care of proper configuration to avoid this state.

Calculation time in such a distributed system is more complex as shown in the previous equation (9).

$$t \quad = t_{rcv} + t_{sched} + t_{start} + t_{calc} + \ t_{persist} + t_{end} + t_{send} \mid m < m_\theta \tag{10}$$

trcv: Time for receiving of the message, which triggers calculation *tsched:* Time required by system to schedule calculation. This is usually not a trivial task to coordinate lifecycle of partitioned calculations of columns in distributed system. All messages must be ordered and when possible, distributed locks shell be avoided. To solve this problem, already named a dedicated Actor Programming Framework was implemented on top of Microsoft Azure Service Bus [17], which provides messaging platform with many features required for this task. In this case message-session is used to ensure that all messages are ordered and dispatched to a single location, where calculation is performed. With this, no distributed locks are possible, and every partition calculation is running on a single thread. Because of this, *tsched* is taken out of algorithm and it remains a part of messaging system.

In this concept, one partition is defined as an Actor, sometimes called *partition Actor*. It hysically owns 1-N columns (as shown in

Figure 1) and it performs calculus over space P_k as defined by equation (4) owned columns only (see Figure 3). This space is much smaller than space λ defined by equation (5).

$$\{p_{k1}, p_{k2}, .., p_{C_{kq}} \quad | \, C_{kq} <= \mathrm{IN} \qquad (11)$$

The Actor Model guarantees that there is only one calculation running on a single column partition across all cores in the cluster. Every partition Actor also holds the potential pool array as defined by the equation (4) and is capable of calculating the algorithm listed in
Figure 2.

The distributed HTM algorithm **SP-Parallel** performs partitioning of Actors inside of the orchestrator node, which plays the role of a scatter operation. Running of calculations in actors on remote nodes is started and awaited on multiple threads inside of the orchestrator. Finally, the Actor model executes actors on nodes in the cluster and results are collected by the orchestrator node, which now plays the role of a gather operation.

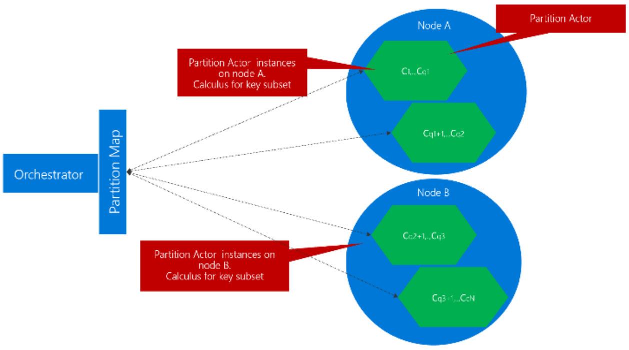

Partition Actor
Node A
Partition Actor instances on node A. Calculus for key subset
C1,...Cq1
Cq1+1,...Cq2
Orchestrator
Partition Map
Node B
Cq2+1,...Cq3
Partition Actor instances on node B. Calculus for key subset
Cq3+1,...CqN

Figure 3 : Partitioned column space. Every partition is implemented as an Actor, which owns subsets of columns from the entire column graph. By providing the number of columns and nodes, the number of partitions can be automatically calculated or explicitly set.

To recap, in this partitioning concept the number of partitions and the number of nodes are known. That means, the SP-Parallel orchestrator code, which initiates placement of partitions must know nodes and can explicitly place a partition to the dedicated node. With this, the Actor model framework can ensure that full calculation is executed as a sticky session on an initiated node. This improves performance and does not require a durable persistence of the calculation state, because the state is kept in the cache.

There is also another approach, which was tested, but it was excluded from this paper. In this (second) approach the orchestrator node does not have any knowledge about the number of nodes. This enables a simpler architecture of the system, but it requires to durably store the calculation state because, after every calculation step, the next step can be initiated on another node. For this reason, nodes must be able to save and load the state to and from durable storage, which adds significant performance costs. The second approach would perform better for shorter calculations with less memory

5. EVALUATION

In this work several experiments have been created, which evaluate the performance of the modified Spatial Pooler algorithm. For all tests MNIST images of 28x28 pixels have been used. First, a single-threaded algorithm was compared against SP-MT (multicore single node SP) on different topologies (results shown for 32x32 columns).

Then the compute time of SP-Parallel was tested for a column topology 200x200 on one, two and three physical nodes. Finally, the performance of several column topologies was tested in a cluster of three physical nodes.

All tests have been executed on nodes with following "*commodity*" configuration on virtual machines in Microsoft Azure cloud:OS: Microsoft Windows Server 2016;

Processor: Intel(R) Xeon(R) CPU E5-2673 v4 @ 2.30GHz, 2295 MHz, 2 Core(s), 4 Logical Processor(s); Physical Memory (RAM):16.0 GB.

The first experiment was designed to demonstrate performance improvements of the SP-MT versus single-threaded algorithm. Remember, both algorithms were running on a single node. As input, MNIST test images with 28x28 pixels were used, and a cortical topology of 32x32 columns. Figure 4 shows the resulting sparse representation of the MNIST image.

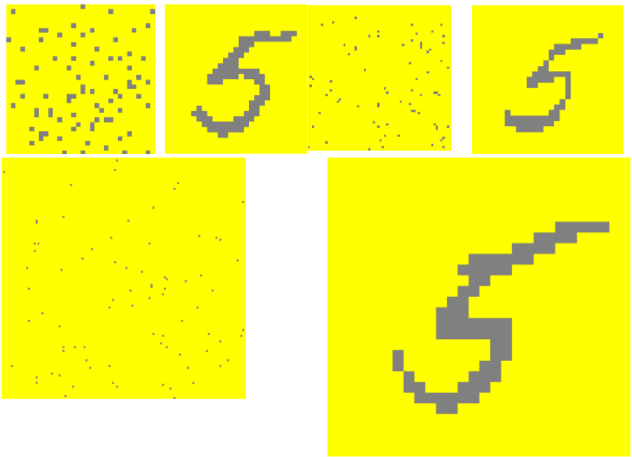


Figure 4: Sparse representations of an MNIST digit with different sparsity in column topologies 32x32 (top-left), 64x64 (top-right) and 128x128 (bottom). As an example, SDR on the top-right with column topology of 64x64 (4096 columns) occupies 2% (81) columns only.

Results shown in Figure 5 indicate that SP-MT is approximately twice faster than SP single-threaded on the indicated VM configuration.

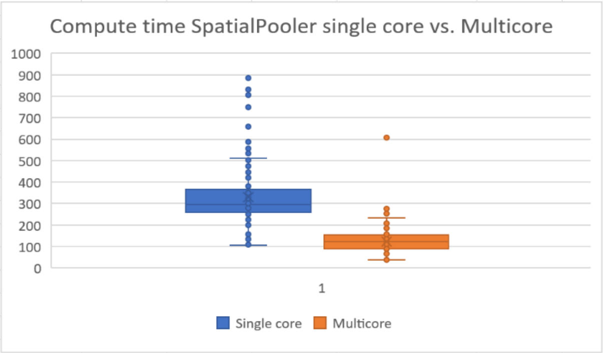


Figure 5 : Performance results, SpatialPooler single-core (SP) versus Spatial Pooler multicore (SP-MT) on a single machine. Tested on Surface Book2 with Microsoft Windows 10 Enterprise, Processor Intel(R) Core(TM) i7-8650U CPU @ 1.90GHz, 2112 MHz, 4 Core(s), 8 Logical Processor(s). MNIST 28x28 test image used 32x32 columns.

In the same experiment, the memory consumption (see
Figure 6) and processing time in milliseconds in dependence of column topology were measured (see Figure 7).

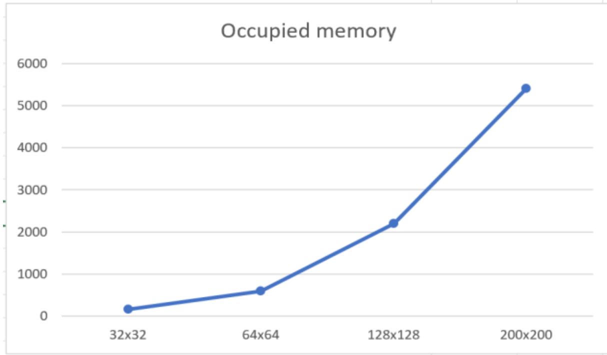


Figure 6 : RAM memory in MB, occupied during execution of the SpatialPooler-MT algorithm on a single node.

By using the same experiment with SP-Parallel instead of SP-MT, topologies with a higher number of columns and multiple nodes were tested. In this experiment learning of the MNIST image was measured on 1, 2 and 3 nodes. As shown in Figure 8 SP-Parallel on a single node needs nearly the same time as SP-MT. This is a good result because it approves that the Actor model framework does not spend significant time on the internal messaging mechanism.

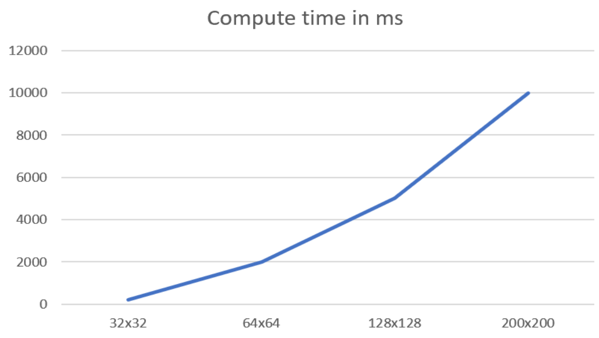


Figure 7 : Spatial Pooler-MT compute time in milliseconds in dependence of column topology.

By adding more nodes to the cluster, performance increases as expected. The optimal number of partitions still must be investigated. As for now, to ensure that calculations on multiple partitions can run in parallel, it should be 2 or 3 times higher than the number of cores on all nodes.

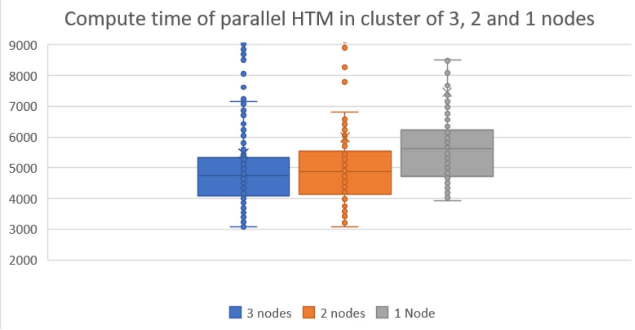


Figure 8 : Learning time in [ms] of a MNIST image 28x28 with 200x200 Columns. Performance on a single node corresponds nearly to performance of SP-MT algorithm.

Figure 9 shows memory and CPU consumption on a single node, while calculation is running on multiple nodes. Independent of column topology, both memory and CPU consumption are shared across nodes in the cluster. As shown by the figure, during initialization time (step 1) memory is increasing, while allocating space for the columns and then it gets stable across the remaining repeating steps 2, 3 and 4 during the iterative learning process.

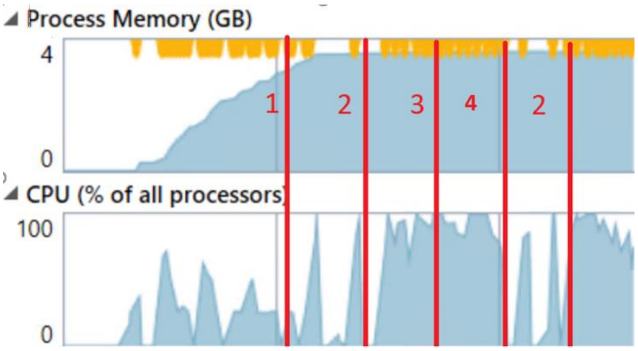


Figure 9 : Process memory on a node while the computation of SP-Parallel is running. At the beginning 1 initialization stage is running, which allocates the required memory. Subsequently, stages 2,3 and 4 are related to overlap, synapse adaption and activation of inactive columns.

Finally, the system was tested to run up to 250000 cortical columns. This configuration allocates 196000000 synapses to sensory input (28x28 sensory neurons) on a proximal dendrite, by used global inhibition of the Spatial Pooler in this experiment.

In this experiment, every column connects to all sensory neurons, which corresponds to a potential connectivity radius of 100%. Topologies 200x200, 300x300 and 500x500 (250000 columns) columns were compared.

Additionally, the initialization time (see
Figure 11) of the Spatial Pooler should not be underestimated. Allocating cortical columns and corresponding synapses takes significantly more time than compute time. Note that compute times for topology 200x200 with 20 and 15 partitions do not indicate significant differences. This is because the number of partitions is higher than the number of cores (3 nodes with 4 logical processors) in both cases.

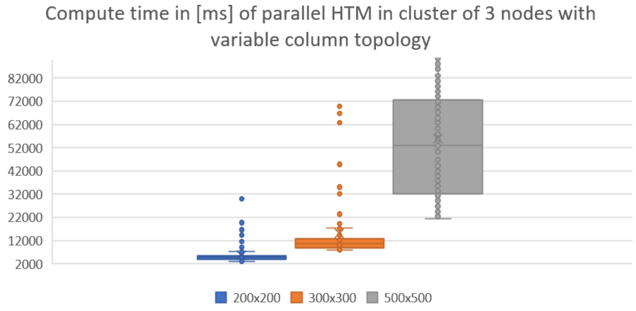


Figure 10 : Compute time of SP-Parallel on three nodes in dependence of column topology.

Having a lower number of partitions than the number of cores would not use available CPU power and having a too high number of partitions would generate too many context switches and slows down the performance.

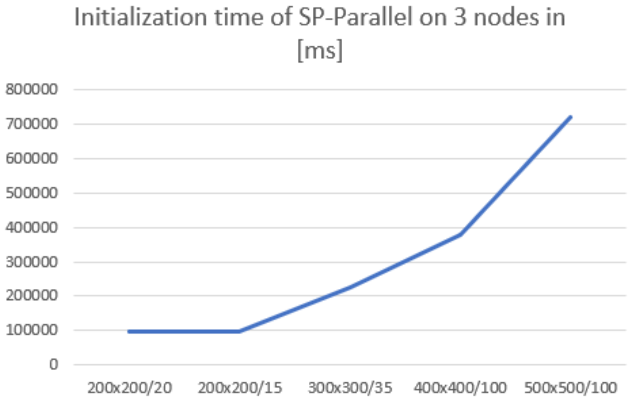


Figure 11 : Initialization time in milliseconds of SP-Parallel in a cluster of 3 nodes in dependence of column topology. Used topologies are 200x200 with 20 partitions, 200x200 with 15 partitions etc.

All presented results were tested with the Actor model implementation, which sticks calculation to specific nodes without state persistence. Persistence of calculations would slow down calculation time. Some additional experiments show (not presented in this paper) that a single column takes approx. 700kb space persisted as JSON. Persisting of partitions described in this paper with thousands of columns would take gigabytes of space and would require a sophisticated algorithm to save and load such state in a very short time. This is one of the future tasks in this context.

6. CONCLUSIONS

Results in this paper show that HTML CLA can efficiently be scaled with an Actor programming model by using higher-level programming languages. Proposed algorithms SP-MT and SP-Parallel can successfully run on multicore and multi-node architectures on commodity hardware, respectively. SP-MT executes on a single node multicore architecture without the Actor Programming Model, which is rather used by SP-Parallel. The modified version of the Spatial Pooler can observe calculations for a high number of cortical columns in the simple Actor model cluster on commodity hardware. The building of algorithms natively in hardware by using lower-level programming languages might show better performance. However, using widely industrial accepted and extremely productive higher-level programming languages enable easier use of compute power of modern cloud environments and enables this technology for use in a wide community of developers. The current version of SP-Parallel and SP-MT rely on the same code base, which will be step by step optimized, for example, in the way how internal sparse representations are implemented, especially when it comes to memory consumption. The goal of this work was to redesign Spatial Pooler for the Actor Programming Model by enabling it for easy horizontal scaling of the multiple nodes. The current implementation supports Windows, Linux and macOS on almost any kind of hardware.

With this approach, cortical regions can be widely distributed across many machines with acceptable costs and performance. Scaling of the Spatial Pooler algorithm is the first step in this research. **Spatial Pooler** produces sparse representations of inputs in the form of active columns. By following findings in neurosciences, generated *sparse representation* can be used as an input for the **Temporal Memory** algorithm. A next step in this research is the design of a highly scalable parallel version of the Temporal Memory algorithm and the design of a cortical network with the used Actor Programming Model approach. Such cortical networks will be capable to build highly interconnected cortical regions distributed in cluster. The high degree of connections should enable powerful sequence learning and more cognitive capabilities.

REFERENCES

[1] Hawkins Jeff, Ahmad, Subutai, Dubinsky Donna, “HTM Cortical Learning Algorithms,” Numenta , 2011. [Online]. Available: https://www.numenta.com/htm-overview/education/HTM_CorticalLearningAlgorithms.pdf.

[2] Hawkins, Jeff and Ahmad, Subutai and Cui, Yuwei , “A Theory of How Columns in the Neocortex Enable Learning the Structure of the World,” Frontiers in Neural Circuits, vol. 11, p. 81, 2017.

[3] Wulfram Gerstner1, Werner M. Kistler2, “Mathematical formulations of Hebbian learning,” US National Library of Medicine National Institutes of Health, vol. 87, no. 5-6, 2002.

[4] J. H. Kaas, “Evolution of columns, modules, and domains in the neocortex of primates,” PNAS, 2012. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3386869/.

[5] “NUPIC,” NUMENTA, [Online]. Available: https://github.com/numenta/nupic.

[6] “HTM.JAVA,” NUMENTA, [Online]. Available: https://github.com/numenta/htm.java.

[7] D. Dobric, “NeoCortexApi,” 2019. [Online]. Available: https://github.com/ddobric/neocortexapi/blob/master/README.md.

[8] Luca A. Finelli, Seth Haney, Maxim Bazhenov, Mark Stopfer, Terrence J. Sejnowski, “Synaptic Learning Rules and Sparse Coding in a Model Sensory System,” PlOS Computational Biology, 2007.

[9] Wulfram Gerstner, Werner M. Kistler, “Mathematical formulations of Hebbian learning,” 2002. [Online]. Available: https://link.springer.com/article/10.1007/s00422-002-0353-y.

[10] M. W. W. Marcin Pietron, “Parallel Implementation of Spatial Pooler in Hierarchical Temporal Memory,” Research Gate, 2016. [Online]. Available: https://www.researchgate.net/publication/301721307_Parallel_Implementation_of_Spatial_Pooler_in_Hierarchical_Temporal_Memory.

[11] C. Hewitt, “Actor Model of Computation,” Cornell University, 2010. [Online]. Available: https://arxiv.org/vc/arxiv/papers/1008/1008.1459v8.pdf.

[12] Microsoft Research, “Virtual Actors,” [Online]. Available: https://www.microsoft.com/en-us/research/project/orleans-virtual-actors/.

[13] Microsoft Corporation, “Service Fabric Reliable Actors,” Microsoft, 2018. [Online]. Available: https://docs.microsoft.com/en-us/azure/service-fabric/service-fabric-reliable-actors-introduction.

[14] Akk.NET, “Akka.NET,” [Online]. Available: https://getakka.net/.

[15] James Mnatzaganian, Ernest Fokoué, Dhireesha Kudithipudi, “A Mathematical Formalization of Hierarchical Temporal Memory's Spatial Pooler,” 2016. [Online]. Available: https://arxiv.org/abs/1601.06116.

[16] M. L. W. H. Mark Pethick, “Parallelization of a Backpropagation Neural Network on a Cluster Computer,” ResearchGate, 2003. [Online]. Available: https://www.researchgate.net/publication/228549385_Parallelization_of_a_Backpropagation_Neural_Network_on_a_Cluster_Computer.

[17] “Cloud Messaging,” Microsoft, [Online]. Available: https://azure.microsoft.com/en-us/services/service-bus/.

AUTHORS

Damir Dobric

CEO and lead software architect of DAENET Corporation specialized in software technologies with strong focus on Cloud Computing, IoT and Machine Learning. Microsoft Regional Director and Microsoft Azure Most Valuable Professional working with Microsoft on helping customers to adopt modern technologies. Teaching software engineering and cloud computing on University of Applied sciences in Frankfurt am Main. Researching at University of Plymouth at the School of Engineering, with the focus on artificial intelligence and distributed systems.

Andreas Pech

Professor of Computer Engineering at Frankfurt University of Applied Sciences (Frankfurt, Germany), Dept. of Computer Science and Engineering, Head of Computational Intelligence Lab. He received his Diplom-Ingenieur degree in Electrical Engineering from TU Darmstadt, Germany in 1986 and his PhD (Dr. rer. nat.) from University of Mayence, Germany in 1990. Previously he worked as a group manager at Deutsche Klinik f Diagnostik, medical research department in Wiesbaden, Germany and as System Engineer at different companies.

Dr. Bogdan Ghita

Associate Professor at the School of Engineering, Computing and MathematicsFaculty of Science and Engineering. Leader of network research within the Centre for Security, Communications, and Network research.

Dr. Thomas Wennekers

A research Associate Professor at the University of Plymouth, UK. He studied Physics in Ulm, Germany, and received his PhD in Computer Science 1998 from the University of Ulm, Germany. His research interests focuses on Hebbian Cell assemblies as dynamic cortical representations in sensory processes and as building blocks of higher cognitive functions.

A 3D SIMULATION FOR THE FEEDBACK LOOP BETWEEN ORBITAL DEBRIS AND FUTURE SPACE ACTIVITIES AND ECONOMY

Qianyu Liu[1], Emmanuel Reyes[2], Yu Sun[3]

[1]El Toro High School, Lake Forest, CA, USA
[2]University of California, Irvine, USA
[3]California State Polytechnic University, Pomona, USA

ABSTRACT

Since the success of SpaceX's reusable launch system program, there has been a massive resurgence in interest in space, hundreds of companies and startups are racing to develop cheaper ways of venturing into the vacuum of space. As a result, the sustainability of the space environment will be put under great danger and pressure, threatening all other future space activities. In the study, we attempt to quantify the chain effect of various forms of space activities and orbital debris using Unity3D, followed by proposing the plan to use NASA's simulation software Orbital Debris Engineering Model (ORDEM) 3.0 and Debris Assessment Software (DAS) 3.0.

KEYWORDS

Orbital Debris, 3D Simulation, Unity3D.

1. INTRODUCTION

The space environment is experiencing the rapidly growing threat of orbital debris, defined by NASA as any object in orbit that does not serve any useful purpose. Examples include nonfunctional spacecraft, spent rocket upper stages, discarded hardware, and fragments from uncontrolled chemical explosions. Debris comes in a wide range of sizes, from less than 1 mm to larger than 40 cm in diameter; orbit, from circular LEO (low Earth orbit) to highly eccentric orbit that reaches beyond GEO (geostationary orbit); and velocity, from less than 3 km/s to 20 km/s. At such high velocity, impact with even minuscule debris can cause devastating damage to any spacecraft, an object with a diameter of 1 cm traveling at a relative speed of 10 km/s contains more kinetic energy than a .50 BMG bullet. As of January 2019, 94% of all catalogued man-made objects in space are debris, over 34,000 are larger than 10 cm, 900,000 are between 1 cm to 10 cm, 128 million from 1 mm to 1 cm, and likely many trillions that are smaller than 1 mm (ESA, 2019).

The danger of orbital debris comes from several factors, unpredictability, detection difficulty, large population, and exponential growth. First, debris is affected by minor forces such as atmospheric drag and gravitational perturbations that will slowly alter their orbit, the resulting orbit is difficult to calculate into the far future, making avoidance maneuvers tighter on time than maneuver with an object whose orbit is known ahead of time. Secondly, due to the tiny sizes of debris, most of them (under 1 cm) cannot be tracked by either ground telescopes nor satellite

David C. Wyld et al. (Eds): SPTM, IPPR, AIS, CSIT, DaMi, AMLA - 2020
pp. 71-78, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100607

detectors, effectively rendering them invisible and impossible to avoid (Watson, 2015). Thirdly, the multitudinous distribution of debris surrounding Earth means that the collision rate is much greater since there are debris in wildly chaotic orbits, coupled with the fact that they are hard to detect and prone to unforeseeable changes, most spacecraft are sitting ducks. Finally, the problem is exacerbated by exponential growth, specifically, this phenomenon is known as the Kessler Syndrome, which describes the cascading growth of debris due to random collisions between the debris themselves. Models have shown that a relatively small amount of debris with very little inclination difference can spread out over time to become a massive band of debris encapsulating the entirety of Earth's celestial sphere (Rex, Eichler, Soppa, Zuschlag & Bade, 1989, p.107-120). Unfortunately, we may have already crossed the point of no return, the amount of debris in orbit has already passed the critical density for uncontrollable growth, meaning that even if we stop putting anything into space, the problem will continue to get worse. The Kessler Syndrome will be a key interest point throughout this paper.

Satellite operation within this dense sphere of debris is incredibly risky, yet in a modern world where society is so dependent on space technology, most are completely oblivious to the fact that the space environment has never been more at risk. The overcrowding of the space environment would greatly hinder human activity and the development of science in space. In this paper, we attempt to quantify these impacts using the Unity3d engine.

We used Unity to replicate a space debris simulation by using game objects to represent debris. The simulation contained a horizontal, vertical and diagonal band of space debris that copied the physics of outer space. Space debris spawns at a random location within their band space and its speed acts accordingly. The closer the debris is to Earth the faster it orbits and the further away it is the slower it orbits. We can see the effects of space debris incrementing by instantiating new debris every time the user left-clicks on their mouse. Each debris contains a box collider that helps us detect the number of close collisions. This number of close collisions can help us determine space activity.

2. RELATED WORK

2.1 Current situation

To understand the full scope of things, it is necessary for this literature review to first introduce the current scale of the problem. In the research compendium Limiting Future Collision Risk to Spacecraft (2013), a number of space organizations examine the complex situation of the future of space debris and its effect on space activities. By studying current data, future technological trends, launch schedules, modeling, and simulations, the current collision rate for a single satellite is around 1 in 1000 per its 15-year life cycle. This means that approximately two operational satellites and three defunct satellites are destroyed every year. However, these supposed collisions have not yet occurred thanks to better tracking technology and avoidance maneuvers.

2.2 Military activities

The earliest predominant purposes were military and national security and it remains so. However, no spacecraft is safe from debris, military ones are no exception, the crowding of the space environment presents both legal and military problems. The United States is more dependent on space than another nation, the threat of orbital debris to military satellites are threats of equivalent seriousness to the national security of the nation. At the same time, countless organizations and groups depend on military space assets to function, many of them integral and irreplaceable to society. There would be incredibly costly for military satellites to be damaged or destroyed (one may cost up to a few billion dollars), but more importantly, the shockwaves of their destruction would be felt by more than just the government, it could be everyone. Due to the nature of communication satellites' transmissions with each other, the loss of a single communication satellite may induce effects similar to that of a full-scale network shutdown. Lieutenant Colonel Joseph S. Imburgia of the USAF asserts that even a short timeout, the loss of communication capabilities could still represent a near-total inability for the nation to defend itself from any form of complications (Imburgia, 2011). In a hypothetical situation where a military satellite is destroyed by space debris but incidentally interpreted as an attack could significantly degrade international relationships and possibly lead to war (Grego, 2011). However, the United States and its allies are deeply in favor and interest to preserve the sustainability of the space environment that their safety is so dependent upon. Military efforts could be made to reduce collision probability or even actively remove debris from orbit, which will surely have influence on other space activities.

2.3 Commercial activities and the space economy

While the military may be the dominant purpose of space activity, space is quickly becoming a commercialized place with unlimited resources and possibilities. An assessment by Jeff Greason and James C. Bennett of the Reason Foundation estimates future commercial activities could become one of the largest industries generating tens of trillions of dollars, a monumental leap from this year's $350 billion (Greason & Bennett, 2019). Examples of commercial activities include satellite communication, space tourism, interplanetary transportation, asteroid mining, etc, all containing immense economic potential. However, space debris has the possibility to shatter all these economic possibilities by making venturing into space overly dangerous for any meaningful commercialization impossible (Weinzierl, 2018). The problem is massively exacerbated by the recent invention of reusable rocket technology, which has lowered launch costs across the board by a dramatic amount (Adrian and Hyman, 2018). This sudden breakthrough has caused major disruption in the aerospace industry, veteran and startup companies alike are rushing to lower the price of venturing into space further, lowering the price tags even more.

This technological success has resulted in a massive resurgence of interest in space, plans that were previously impossible are now within reasonable reaches. Due to both demands can capability, reusable rockets are now planned to be used to build satellite mega-constellations for the purpose of worldwide ultra-fast internet connection, such as the 600 from OneWeb and the 12,000 from SpaceX. If their plans do indeed come to reality, it would be one of the biggest technological achievements; while at the same time, the greatest catastrophe to the space environment. As the sheer size of the fleets would increase the already-congested orbit by more than six times, inevitably causing collision rates to skyrocket. (Virgili et al, 2018) (Le May, Gehly, Carter, and Flegel, 2018). Along the way, both the space economy and environment will certainly be affected, the extent of effect are also sufficiently understood, however, this paper will focus partly on the cross-effect between the two instead.

2.4 Scientific activities

Due to its unique environment, space is an important place for science, many satellites have been launched into space for various purposes. Today, humans use many items that were invented thanks to orbital sciences without knowing it, and clearly, we have a dependence on it. However, research shows that orbital debris is posing an increasingly large threat to all orbital scientific operations. As the largest spacecraft ever constructed, the International Space Station had performed a total of 25 debris avoidance maneuvers since 1999, while sustaining countless micro-impacts. In the future with an unmanaged debris environment, the orbital laboratory may no longer be able to avoid debris due to the extremely population and render uninhabitable, effectively abandoning one of the most prolific science labs we have (Johnson & Klinkrad, 2009). A slow-down in scientific advancement would directly hinder the progression of space-related technology. Other scientific spacecraft are also affected. Space telescopes, crucial for the purpose of astronomy as their location allows them to avoid atmospheric turbulence, light pollution, and various other factors that affect observation quality (Hotz, 2017). Earth observation satellites, which are heavily relied upon by environmental scientists and meteorologists to gather information on the surface of Earth, with considerable usage in a wide variety of sciences (Durrieu & Nelson, 2013). The limitations to experiment and test things in space will certainly induce a stagnation in the development of space-related technology, negatively affecting the advancement of astronautics.

3. SOLUTION AND METHODOLOGY

3.1 Overview of the Solution

In order to simulate and visualize the orbital debris effects, we have built a 3D simulation environment using the game engine Unity3D. We simplify the object 3D models by ignoring the modeling details, but the core parameters about the moving states of the debris, which allows to customize the simulation requirements.

3.2 Components

Unity is a game development engine that can be used to create simulations and games in two or three dimensions. In our case we used Unity to replicate a Space Debris Simulation. In our model, each color block represents an object in space. The red block represents a space debris that orbits vertically around the white cube which represents the Earth. The blue block represents a space debris that orbits on the horizontal axis and the pink block represents a space debris that orbits diagonally around the white block. In order to see more space debris, spawn the user must left click on their mouse. Depending on where these new space debris spawn their speed will act accordingly. The user can add as many space debris as they would like to see the effects of increasing debris. Every time the user left-clicks a new space debris is instantiated on every band.

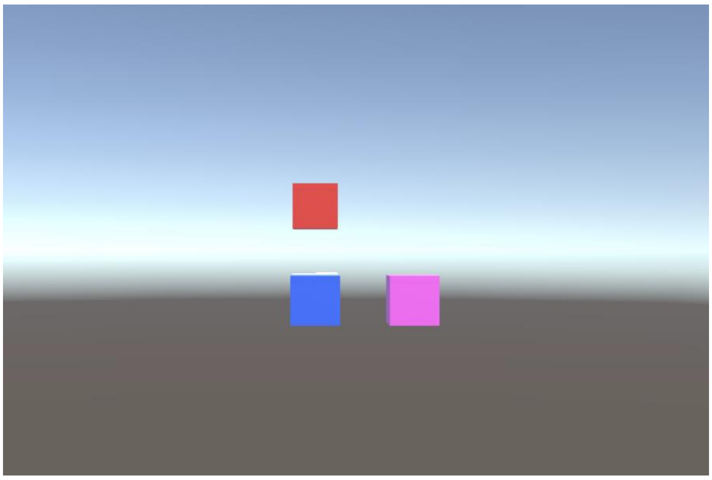


Figure 1. The initial state of the 3D simulation.

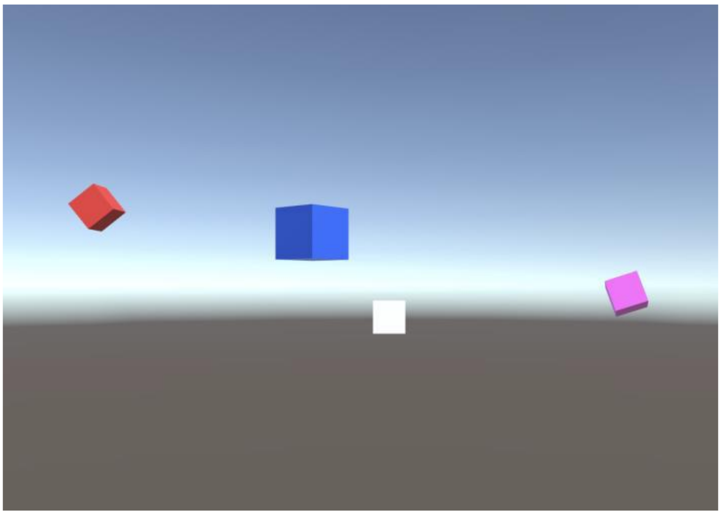


Figure 2. The moving state of the 3D simulation.

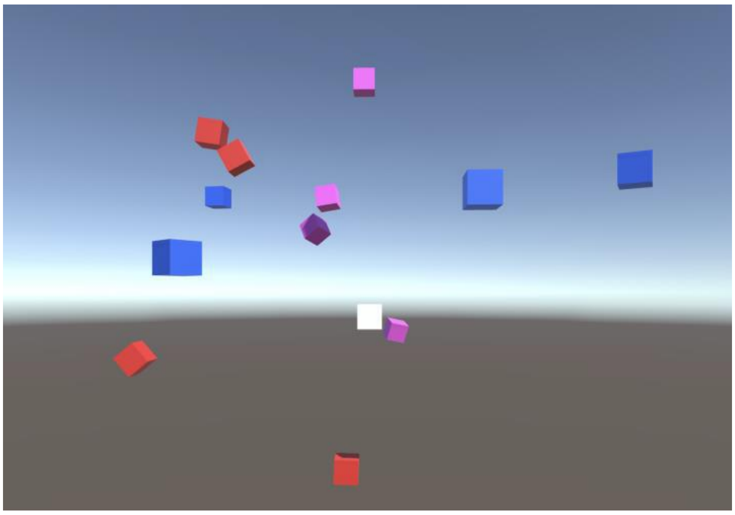
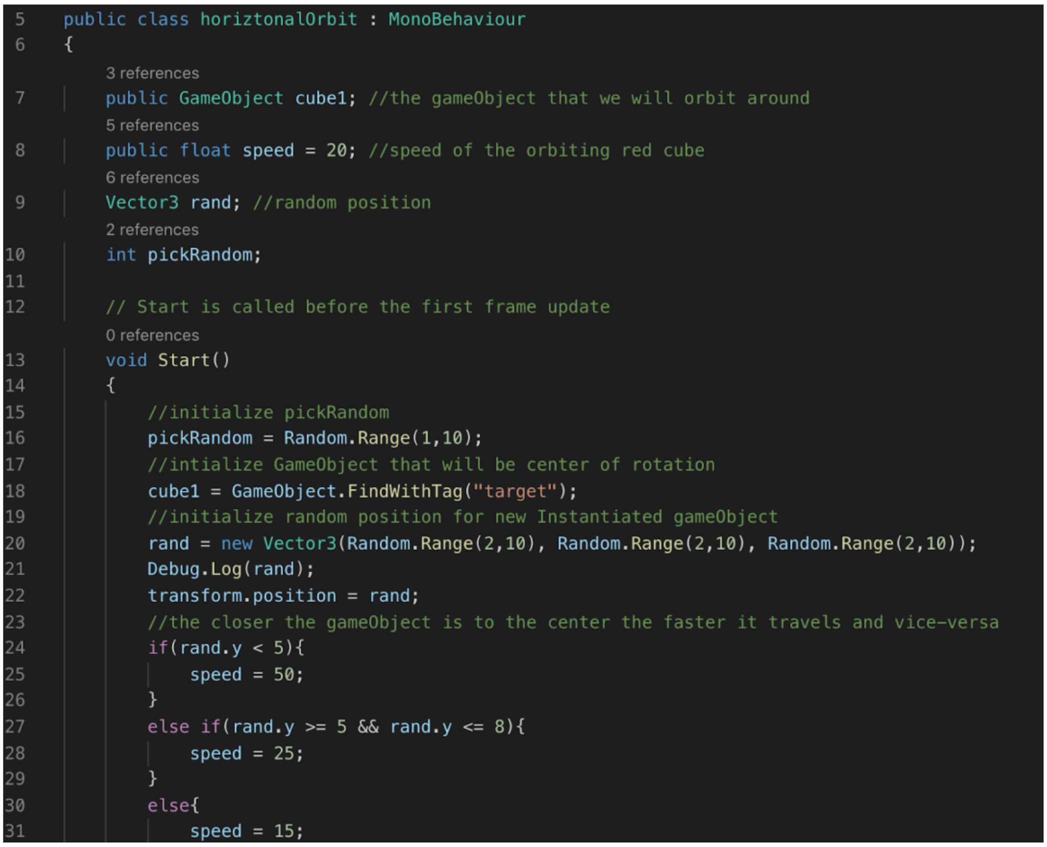


Figure 3. The increased scale of moving state of the 3D simulation.

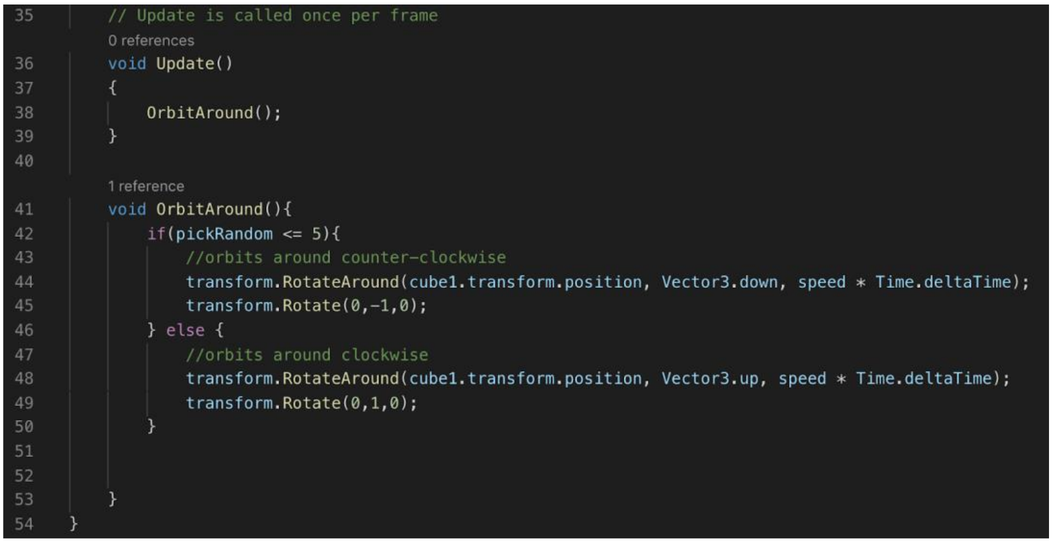
Here are some screenshots of the code that makes this space debris simulation work. If you read the comments on the code it helps explain what each line of code does. Here, we made sure that our space debris orbits around the centered block which represents the Earth and is labeled “cube1.” We then initialize its default speed and in the “Start” function we set a random location for our debris to spawn within its specified band range. Next, we set the speed of the debris according to its y position. Every time a new space debris is instantiated, we then randomize its orbit. If the random number that is picked is less than or equal to five than the debris will orbit counterclockwise otherwise it will orbit clockwise.



```
public class horiztonalOrbit : MonoBehaviour
{
    3 references
    public GameObject cube1; //the gameObject that we will orbit around
    5 references
    public float speed = 20; //speed of the orbiting red cube
    6 references
    Vector3 rand; //random position
    2 references
    int pickRandom;

    // Start is called before the first frame update
    0 references
    void Start()
    {
        //initialize pickRandom
        pickRandom = Random.Range(1,10);
        //intialize GameObject that will be center of rotation
        cube1 = GameObject.FindWithTag("target");
        //initialize random position for new Instantiated gameObject
        rand = new Vector3(Random.Range(2,10), Random.Range(2,10), Random.Range(2,10));
        Debug.Log(rand);
        transform.position = rand;
        //the closer the gameObject is to the center the faster it travels and vice-versa
        if(rand.y < 5){
            speed = 50;
        }
        else if(rand.y >= 5 && rand.y <= 8){
            speed = 25;
        }
        else{
            speed = 15;
```

Figure 4. An excerpt of the Orbit controller code.



```
    // Update is called once per frame
    0 references
    void Update()
    {
        OrbitAround();
    }

    1 reference
    void OrbitAround(){
        if(pickRandom <= 5){
            //orbits around counter-clockwise
            transform.RotateAround(cube1.transform.position, Vector3.down, speed * Time.deltaTime);
            transform.Rotate(0,-1,0);
        } else {
            //orbits around clockwise
            transform.RotateAround(cube1.transform.position, Vector3.up, speed * Time.deltaTime);
            transform.Rotate(0,1,0);
        }


    }
}
```

Figure 5. An excerpt of the Orbit rotation code.

4. FUTURE WORK

The approach we used were based purely on 3D modeling and simulation from the models we have created from scratch. To get a better result, we have recently gained access to a model and simulation for the debris population growth provided by NASA's Orbital Debris Engineering Model 3.1 (ORDEM 3.1) and Debris Assessment Software 3 (DAS 3.0) model. For the future work, four runs will be conducted on both software, each with its specific setting to simulate the following:

- No spacecraft, no launches, debris only
- Business as usual (spacecrafts and launches)
- Launches with mild effort to mitigate damage and reduce debris
- Launches with significant effort to mitigate damage and reduce debris

Of each individual run, the following data will be gathered:

- Number of collisions
- The probability of collisions
- The number of close encounters
- Final debris count

Data will be assessed to estimate the collision probability for debris of other sizes, then the entirety of collision probability will be combined to calculate the total probability of collision over time. The total probability will then be used to qualitatively and quantitatively estimate the effect of orbital debris on spacecrafts and other debris alike.

5. CONCLUSIONS

The current situation regarding orbital debris is no doubt concerning, but nonetheless within our technological capability to mitigate. However, with the fast-approaching future of 5G network and global internet coverage, the usage of space will become precious and crowded. As demonstrated by our research and others prior, the overcrowding of the space environment will be disastrous to society in almost every way possible, and incredibly difficult so manage once the population reaches a critical number. The dynamic space environment and unpredictability of human actions make this issue all the more difficult. But the problem is not too late to manage, with extensive monitoring of rocket bodies and effective post-mission disposal by satellite companies will ensure a sustainable orbital environment for the future society.

REFERENCES

[1] Chen, S. (2011). The Space Debris Problem, Asian Perspectives. Retrieved from: https://search.proquest.com/docview/928083848?accountid=160992

[2] ESA, (2019), Space Debris by the Numbers. Retrieved from: https://www.esa.int/Safety_Security/Space_Debris/Space_debris_by_the_numbers

[3] Watson, T., (2015), Incoming Space Junk: A Scientific Opportunity, Nature. Retrieved from: https://www.nature.com/news/incoming-space-junk-a-scientific-opportunity-1.18642

[4] Rex, D., Eichler, P., Soppa, U., Zuschlag, J., and Bade, A., (1989), Space Debris - Origin, Evolution and Collision Mechanics, G. W. Heath (Ed.), San Diego, CA; American Astronautical Society.

[5] National Academies of Sciences, Engineering, and Medicine, (2011), Limiting Future Collision Risk to Spacecraft. Retrieved from: https://www.nap.edu/catalog/13244/limiting-future-collision-risk-to-spacecraft-an-assessment-of-nasas

[6] S. Imburgia, J. (2011), Space Debris and Its Threat to National Security: A Proposal for a Binding International Agreement to Clean Up the Junk. USAF. Retrieved from: https://wp0.vanderbilt.edu/wp-content/uploads/sites/78/Imburgia-FINAL-CR-pdf.pdf

[7] Grego, L. (2011). Security in Space: What is at Stake and How do we Move Forward? John Hopkins University. Retrieved from: https://search.proquest.com/docview/928083843?accountid=160992

[8] Greason, J., C. Bennett, J., (2019). The Economics of Space: An Industry Ready to Launch, Reason Foundation. Retrieved from: https://reason.org/wp-content/uploads/economics-of-space.pdf

[9] Weinzierl, M., (2018). Space, the Final Economic Frontier, Journal of Economic Perspective. Retrieved from: https://pubs.aeaweb.org/doi/pdf/10.1257/jep.32.2.173

[10] Adrian, H., Hyman, W., (2018), Reusable Launch System: The Gateway to The Future of Space Travel. Retrieved from: http://www.pitt.edu/~budny/papers/8226.pdf

[11] B.Virgili, B.; Dolado, J.; Lewis, H.; Radtke, J.; Krag, H.; Revelin, B.; Cazaux, C.; Colombo, C.; Crowther, R.; Metz, M., (2016), Risk to Space Sustainability From Large Constellations of Satellites,Acta Astronaut. Retrieved from: https://www.sciencedirect.com/science/article/abs/pii/S0094576516300820?via%3Dihub

[12] L. May, S., Gehly, S., A. Carter, B., Flegel, S., (2018), Space Debris Collision Probability Analysis For Proposed Global Broadband Constellations, Acta Astronautica. Retrieved from: https://www.sciencedirect.com/science/article/abs/pii/S0094576518304375

[13] Johnson, N., Klinkrad, H.. (2009). The International Space Station and the Space Debris Environment: 10 Years On, ResearchGate. Retrieved from: https://www.researchgate.net/publication/228519499_The_International_Space_Station_and_the_Space_Debris_Environment_10_Years_On

[14] L. Hotz, H., (2017), A Band Of Junk Clutters Space --- Debris Imperils Hubble Space Telescope and Equipment Used for Phones, Security, Weather, The Wall Street Journal. Retrieved from https://search.proquest.com/docview/1938015419?accountid=160992

[15] Durrieu, S., F. Nelson, R., (2013), Earth Observation From Space-The Issue of Environmental Sustainability, Space Policy, Retrieved from: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20140011102.pdf

REBD: A CONCEPTUAL FRAMEWORK FOR BIG DATA REQUIREMENTS ENGINEERING

Sandhya Rani Kourla, Eesha Putti and Mina Maleki

Department of Electrical and Computer Engineering and Computer Science,
University of Detroit Mercy, Detroit, MI, 48221, USA

ABSTRACT

Requirements engineering (RE), as a part of the project development life cycle, has increasingly been recognized as the key to ensuring on-time, on-budget, and goal-based delivery of software projects;compromising this vital phase is nothing but project failures. RE of big data projects is even more crucial because of the main characteristics of big data, including high volume, velocity, and variety. As the traditional RE methods and tools are user-centric rather than data-centric, employing these methodologies is insufficient to fulfill the RE processes for big data projects. Because of the importance of RE and limitations of traditional RE methodologies in the context of big data software projects, in this paper, a big data requirements engineering framework, named REBD, has been proposed. This conceptual framework describes the systematic plan to carry out big data projects starting from requirements engineering to the development, assuring successful execution, and increased productivity of the big data projects.

KEYWORDS

Big data, requirements engineering, requirements elicitation, knowledge discovery.

1. INTRODUCTION

Requirements engineering (RE) is the first and most critical phase of the Software Development Life Cycle (SDLC). RE is the branch of engineering concerned with the real-world goals for, functions of, constraints on systems, and the relationship of these factors to precise specifications of system behavior [1-4]. Software requirements engineer (SRE) is responsible for translating stakeholders' needs, desires, and wishes of a software system into precise and formal software requirements specification. SREs need to communicate effectively and frequently with stakeholders to elicit, analyze, model, and manage their requirements. Doing this at the early stage of the software development helps ensure requirements are meeting stakeholders' needs while addressing compliance and staying on schedule and within budget [3].

In contrast, it has been indicated in many researches that performing improper and careless RE is one of the main sources of the time-consuming rework, inadequate deliveries, budget overruns, and consequently failures of the projects [3, 5-7]. The HMS Titanic (1912), the Mars Climate Orbiter (1999), the Apollo 13 (1970), Space Shuttle Challenger (1986), and Space Shuttle Columbia (2002) projects are some examples of high-profile projects broke down due to poor requirements engineering [7]. Also, according to PMI's Pulse of the profession (2018), 35% of companies proclaimed that the primary reason for their project failures is inaccurate requirements gathering [8]. As a consequence, performing a proper RE at the early stages of project development is critical in the success of projects, especially for big data projects.

David C. Wyld et al. (Eds): SPTM, IPPR, AIS, CSIT, DaMi, AMLA - 2020
pp. 79-87, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100608

In a newly released study, International Data Corporation (IDC) forecasts that the big data technology and analytics software market, which in 2018 reached $60.7 billion worldwide, is expected to grow at a five-year compound annual growth rate (CAGR) of 12.5% [9]. Big data is a technology that deals with data sets that are too large or complex to handle with traditional data processing techniques for capturing, storing, analyzing, searching, sharing, transferring, visualizing, querying, and updating of data. The main characteristics of big data technologies are 5V's: volume, velocity, variety, volatility, and variability [4, 10-12]. Volume refers to the massive amount of data; Velocity refers to the high growth rate of incoming data that needs to be processed and analyzed; Variety refers to many different forms of data; Volatility refers to the duration which data is valid and should be stored in the repository; Variability refers to data whose context changes invariably. For big data projects, it is essential to define the targets of the project at earlier stages. Time and cost reduction, finding insights from statistics, and optimizing selection making are the most common goals that make an organization to appreciate big data projects [4, 11]. RE phase in big data projects' development helps in achieving these goals and finding the business values associated with projects; These values help stakeholders to understand the importance of the project and its value in the market [13].

However, employing traditional RE activities, including gathering, analyzing, modeling, validating, and documenting requirements, is not quite efficient for big data projects. The main reason is that traditional RE methods are user-centric and deals with requirements apparent to users, while big data RE methodologies should be data-centric as well. It means there are lots of potential information, hidden patterns, and knowledge that can be extracted and discovered from a large amount of historical and actual data existing in big data projects. Moreover, RE activities in big data should isolate the requirements for infrastructures, analytical tools and techniques, and end-user applications using big data [14]. This isolation is needed because gathering requirements and defining all 5V's characteristics for end-user applications is challenging [12, 15], and SREs are unfamiliar about performing this task.

In this paper, we have proposed a big data requirements engineering framework, named REBD because of the importance of RE in the software project's success and the limitation of traditional RE activities for big data projects. This framework explains the detailed steps for carrying out RE activities on different software projects development for improving their success rate. This framework first identifies the project type and then performs corresponding RE processes on it.The proposed framework also helps to eradicate many challenges regarding knowledge discovery and balancing both data infrastructure and software development (SD) in big data projects.

The rest of the paper is organized as follows. After reviewing related works in section 2, section 3 describes traditional requirements engineering activities. Section 4 briefs big data requirements engineering; Section 5 explains the proposed REBD framework, and section 5 wraps up the research.

2. RELATED WORK

Requirements engineering in the context of big data applications is a hot research topic that attracted the attention of researchers in recent years. An analysis of the state of the art of big data RE research studies shows that little research has been conducted in this area by 2018 [11]. The investigation areas of this research included the phases of the RE process, type of requirements, application domains, RE research challenges, and solution proposals intended by RE research in the context of big data applications. In the following, some of the related work for big data RE will be presented.

To understand the context of the problem for the big data software applications, an empirically derived RE artifact model has been proposed in [15], which is equivalent to domain models. The proposed model can capture the fundamental RE components and the relationships concerning the development of big data software applications. “How can the requirements of a project be identified and analyzed to determine the suitability of a combined big data technology application?”. To answer this question, a new process model has been proposed, detailed, and evaluated in [1]. This model considers the compound requirements, Knowledge Discovery in Databases (KDD), and a big data classification framework to identify the most relevant requirements of big data projects. Identified requirements for building big data applications must address big data characteristics (5V’s) in terms of quality requirements. Quality requirements, also known as quality attributes, are those functional or nonfunctional requirements that used to measure the system’s performance, such as reliability and availability. A new approach is proposed in [10] to ensure that the big data characteristics have adequately addressed in the specification of quality requirements.

Even though there are many requirements elicitation techniques in traditional RE, the researchers believe that more efficient tools and techniques should be employed to identify all requirements, business values, and knowledge from big data. Machine learning, deep learning, natural language processing, and data mining are some data analysis technologies that can use to discover requirements and valuable knowledge from big data [16-18]. Also, actionable use case (AUC) diagram is one of the efficient tools introduced in [13] that allows identified business values unleashed from data to be depicted in the data scientist’s model, together with their roles in the software and their interactions with other software components. The AUC diagrams enhance the users’ experience, optimize the systems’ utility, and consequently maximize profit. Process mining, described in [19], is another efficient method that helps SRE to elicit, prioritize, and validate requirements from big data using execution logs, process discovery, and conformance techniques. The capability of process mining to discover valuable insights from event logs and processes of the system helps SRE to eradicate many challenges of traditional RE.

3. TRADITIONAL REQUIREMENTS ENGINEERING

As mentioned earlier, RE is one of the essential phases of the project’s life cycle, and the failure of the RE leads to the failure of the project. So, software requirement engineers (SRE) are responsible for conducting the RE activities in a very well-mannered way, which can resolve many conflicts between stakeholders as well as among requirements. The main activities of traditional RE are listed below [2-3,13]:

A. Requirements Elicitation: Requirement elicitation as a critical activity in the requirement development process is the practice of uncovering and gathering the stakeholders’ needs and desires from the system in terms of functional and nonfunctional requirements. To have a productive elicitation and dig all of the stakeholders’ requirements, SRE should be able to communicate effectively with all stakeholders. Multiple techniques such as brainstorming, card sorting, laddering, interviews, prototyping, domain analysis, Joint Application Development (JAD), and Quality Function Deployment (QFD) can be employed to conduct requirements elicitation.

B. Requirements Analysis: Requirement analysis is the practice of defining the system boundaries and analyzing elicited and documented requirements to make sure that they are clear, concise, understandable, unambiguous, consistent, and complete. Also, requirements agreement, which is a process of resolving the conflicts in the requirements derived from different sources, is part of this activity. Use cases and user stories are often useful tools for this purpose.

C. Requirements Modeling: The requirement modeling intends to mold raw requirements into technical representations of the system. This activity guides SRE to archive all requirements during the specification process. Proper requirements representation eases the communication of requirements and conversion of those requirements into system architecture and design. To represent the analyzed functional and nonfunctional requirements in any systems, various modeling techniques can be employed, such as use cases, user stories, natural languages, formal languages, and a variety of formal, semiformal, and informal diagrams [3].

D. Requirements Validation: Requirements validation is the practice of checking and revisingrequirements to ensure that the specified requirements meet customer needs. Requirements validation also involves checking that (a) the system does not provide the functions that the customer does not need, (b) there are no requirements conflicts, and (c) time and budget constraints will meet. Systematic manual analysis of the requirements, test-case generation, comparative product evaluation tools, or some of the requirements elicitation techniques can be used for requirements validation [3].

E. Requirements Specification: It is the process of documenting the stakeholders' needs in terms of precise and formal software requirements specification. A software requirement specification (SRS) is a document that uses as a foundation of software development and acts as a contract between the software stakeholders and the developers. As a consequence, preparing an accurate SRS report at the end of the RE phase is critical in the success of projects. SRS contains the mission, scope, goals of the project, software and hardware requirements of the project, as well as functional and nonfunctional system requirements in terms of formal, semiformal, and formal models created in the previous processes.

4. BIG DATA REQUIRMENTS ENGINEERING

As mentioned earlier, the traditional RE methods are user-centric; focus on requirements apparent to users. However, big data RE methodologies should be data-centric as well because there are lots of potential information and hidden patterns that can be extracted from data by data scientists. As a consequence, big data RE model proposed in [13] consists of three different types of processes: processes drove by software requirements engineer (SRE), processes drove by data scientists (DS), and processes drove jointly by software requirements engineers and data scientists (SRE/DS). Table 1 demonstrates the RE activities for the big data products, including the RE activities, the responsible person for executing each activity, and their artifacts. Moreover, the column "Included in" indicates whether each activity is carried out in the traditional RE model (TRE) and/or big data RE model (BDRE).

From the table, it is clear that compared to the traditional RE, big data RE contains a few extra steps, including:

- **Data Acquisition:** This activity can be explained as collecting, gauzing, and fine-tuning the data before storing it in the data repository [13, 20]. The traditional collecting methods cannot make data acquisition because of the characteristics of big data and the high cost associated with it. Hence, data scientists use different data acquisition methods to discard useless data and keep important ones [21].

- **Data Analysis and Value Discovery:** In this activity, the acquired big data is analyzed by data scientists to reveal information and discover business values. Since data analysis is a time taking task and failing in better analysis will lead to consider a wrong decision, different data analysis technologies such as machine learning, deep learning, natural language processing, data mining, text analytics and predictive analytics help in fixing this issue [13, 16-18]. Moreover, techniques like process mining [19] and tools like smart grids [22] can be employed to accelerate the process of knowledge discovery from big data.

- **Use Case Consolidation:** Consolidation is merging the work and building models by software requirements engineers and data scientists. This consolidation is represented as an actionable use case diagram. The AUC diagram for big data software proposed in [13] is a merging of the data scientist model and the traditional use case. Once the values are defined and presented, the consolidated model is ready to be presented and discussed with the customers.

In addition to the requirements matrices, the big data SRS report contains business values extracted from data and AUC diagrams. Moreover, to ensure that big data characteristics are appropriately addressed, SRS should contain the quality attribute matrices. The quality attribute matrix is designed by intersecting a big data characteristic with a quality attribute and then identifying the system's quality requirements that apply to that intersection during big data RE process, as explained in [10].

5. CONCEPTUAL REBD FRAMEWORK

In this section, we describe our proposed big data requirements engineering framework, REBD, as depicted in Figure 1. This framework explains the planned approach for carrying out big data projects for improving the success rate. The purpose of this conceptual framework is to eradicate challenges like deciding whether to perform big data requirements engineering or traditional requirements engineering, knowledge discovery, and balancing the efforts to have the successful execution as described in the following.

Table 1. A Requirement Engineering Model for Big Data Software

Process	**Performed by**	**Output**	**Included in**
Requirements Elicitation	SRE/DS along with Customers	List of requirements	TRE and BDRE
Data Acquisition	DS	Filtered data used for knowledge discovery	BDRE
Requirements Analysis	SRE	Technical specifications and SRE models	TRE and BDRE
Data Analysis	SRE	Extracted values and DS model	BDRE
Use case Consolidation	SRE /DS	Combined SRE and DS model	BDRE
Requirements Modelling	SRE	Finalized AUC diagrams	TRE and BDRE
Requirements Validation	SRE	Validated requirements	TRE and BDRE
Requirements Specification	SRE	SRS report	TRE and BDRE

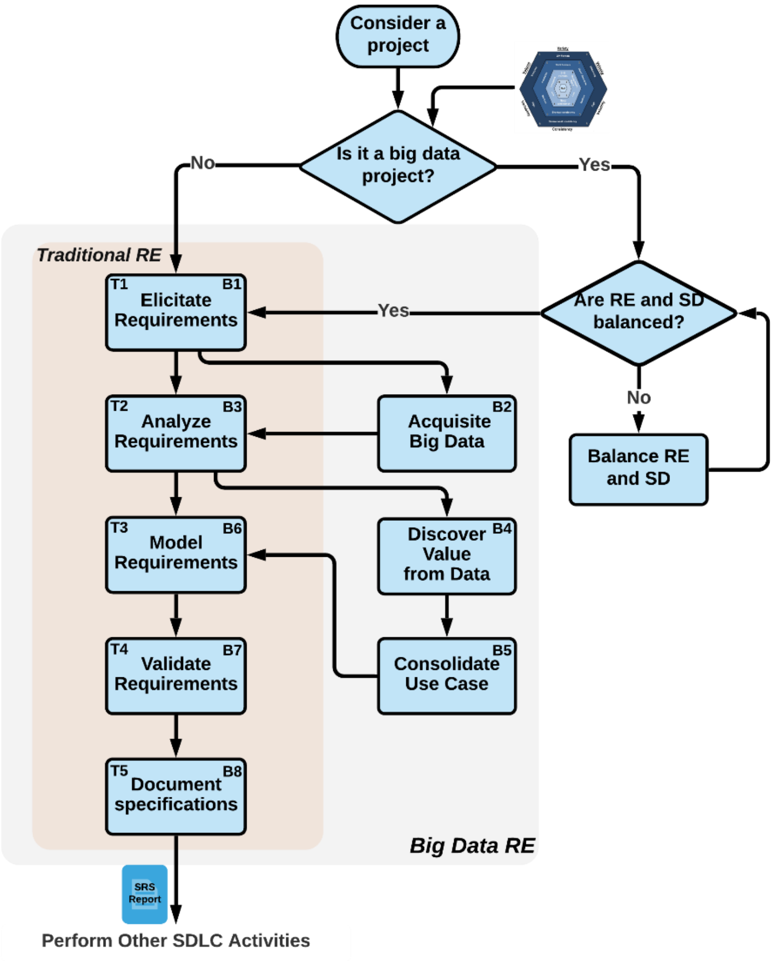

Consider a project
Is it a big data project?
No
Yes
Traditional RE
T1 B1 Elicitate Requirements
Yes
Are RE and SD balanced?
No
Balance RE and SD
T2 B3 Analyze Requirements
B2 Acquisite Big Data
T3 B6 Model Requirements
B4 Discover Value from Data
T4 B7 Validate Requirements
B5 Consolidate Use Case
T5 B8 Document specifications
Big Data RE
SRS Report
Perform Other SDLC Activities

Figure 1. Requirements Engineering for Big Data (REBD) framework

Initially, when the project starts to develop, it is crucial to derive and analyze its characteristics to find whether the project falls under the category of big data or traditional projects. Doing this at the earliest phase of the project will save a considerable amount of time spent on the project lifecycle. To identify the category of the project, we have to figure out whether the recognized characteristics are big data characteristics or not. As mentioned earlier, the main characteristics of big data projects are 5V's: volume, velocity, variety, volatility, and variability. Next, a comprehensive quantitative classification model proposed by [1] is used to verify the project type. This model combines the recognized big data characteristics to figure out the project type. As shown in Figure 2, this classification model contains five layers in which the top four layers are willful to give an exact value depends on the severity, while the fifth level (NULL) is considered for characteristics that cannot be determined presently. If one of the supporting characteristics is assigned to this level, these will not be further utilized. Finally, all assigned values are added up and divided by the number of the addressed characteristics not assigned to the NULL layer. If the calculated assessment value is greater than or equal to 1.33, an application considers being a big data project.

If the project considers being in the group of traditional projects, traditional RE activities should carry out in the order specified in Figure 1 by T1 to T5 labels. However, if the project is found to be in the group of big data projects, then it should be checked whether RE and SD is balanced or not.

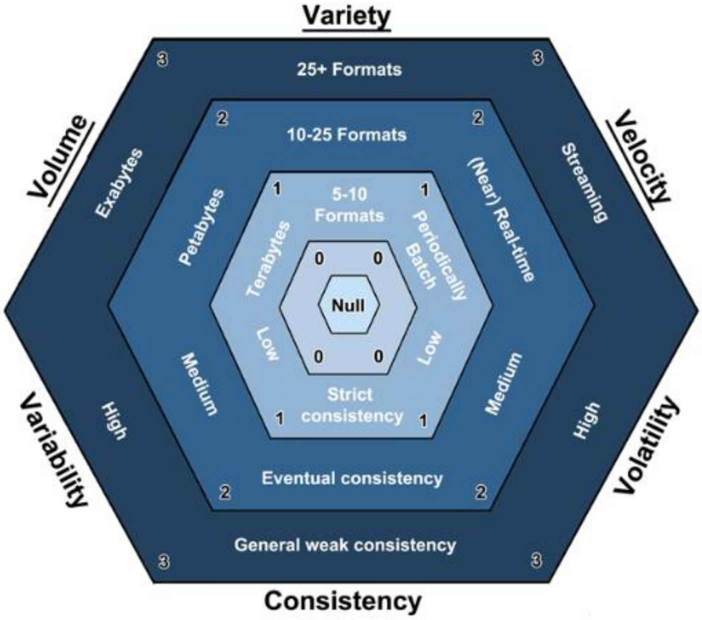


Figure 2. Big data characteristics and classification model [1]

Balancing RE and SD determines the challenges related to RE and SD activities. Challenges related to RE is identifying 5V's, which has been done already, and the challenges associated with SD are designing, coding, testing, and maintaining a big data end-user application. SD challenges are very much concerned because it may lead to the failure of the project, even after spending a lot of time in RE activities. To resolve this issue, a model has been proposed by [14], which contains a separate unit for doing the research needed for RE and SD. Obviously, anything without proof of concepts (PoC) and research may cause a shortage of time. The research and PoC, which give the optimal solution, should be used in the project practice. This ensures the balance between RE and SD, and the development of novel and better big data end-user applications. Once there is a solution for all challenges, then big data RE activities should carry out in order specified in Figure 1 by B1 to B8 labels.

After gathering engineered requirements documented in the SRS report, other processes of the software development life cycles, including design, implementation, testing, deployment, and maintenance, will start to produce software with the highest quality and lowest cost in the shortest time possible.

6. CONCLUSION

One of the main reasons behind the failure of many projects is improper requirements engineering and failing to capture the stakeholders' requirements, which leads to time-consuming rework, inadequate deliveries, and budget overruns. So, performing requirements engineering is a crucial phase of the project's development lifecycle. As big data is one of the new and emerging technology, developing big data projects and improving their success rates is an excellent achievement for companies. However, employing traditional RE activities is not sufficient to guarantee big data projects' development success because of the 5V's characteristics of big data.There is a lot of knowledge, valuable information, business values, and quality features hidden in a large amount of historical and actual raw data existing in big data projects. To analyze thesedata and discover hidden requirements, data scientists need to collaborate with SRE in big data RE activities.

In this paper, a big data requirements engineering framework, called REBD, has been introduced by considering the importance of RE as well as limitations of traditional RE methods. This framework explains the detailed steps for carrying out RE activities on different software projects and can be used to increase the success rate of project development. The proposed framework also considers the importance of balancing RE and SD in big data projects.

Although the current proposed framework is assumed to eradicate the challenges linked with big data, RE and SD, it should be tested practically for the project development to validate its efficiency. This research will be extended to discover better tools and techniques for knowledge discovery from big data. Also, the management of big data quality requirements in a preferred way will be investigated in the future.

REFERENCES

[1] M. Volk, N. Jamous, and K. Turowski, "Ask the right questions: requirements engineering for the execution of big data projects," presented at the 23rd Americas Conference on Information Systems, SIGITPROJMGMT, Boston, MA, Aug. 2017, pp 1-10.

[2] J. Dick, E. Hull, and K. Jackson, *Requirements Engineering,* 4th edition, Springer International Publishing, 2017.

[3] P.A. Laplante, *Requirements Engineering for Software and Systems,* 3rd Edition. Auerbach Publications (T&F), 20171024. VitalBook file, 2017.

[4] D. Arruda, "Requirements engineering in the context of big data applications," ACM SIGSOFT Software Engineering Notes, 43(1): 1-6, Mar. 2018.

[5] A.G. Khan, et. al., "Does software requirement elicitation and personality make any relation?" Journal of Advanced Research in Dynamical and Control Systems. 11. 1162-1168, 2019.

[6] A. Hussain, E. Mkpojiogu, and E. Kamal, "The role of requirements in the success or failure of software projects," presented at the International Soft Science Conference (ISSC), Malaysia, 2016, pp 6-7.

[7] T.A. Bahill and S.J. Henderson, "Requirements development, verification, and validation exhibited in famous failures," Systems Engineering, 8(1): 1–14, 2005.

[8] Pulse of the Profession 2018: Success in Disruptive Times, 2018.

[9] C. Gopal, et al., "Worldwide big data and analytics software forecast, 2019–2023," IDC Market Analysis, US44803719, Sept. 2019.

[10] I. Noorwali, D. Arruda, and N. H. Madhavji, "Understanding quality requirements in the context of big data systems," presented at the 2nd International Workshop on Big Data Software Engineering (BIGDSE), Austin, USA, May 2016, pp. 76-79.

[11] D. Arruda and N.H. Madhavji, "State of requirements engineering research in the context of big data applications," presented at the 24th International Working Conference on Requirements Engineering: Foundation for Software Quality (REFSQ), Utrecht, The Netherlands, Mar. 2018, pp 307-323.

[12] M. Volk, D. Staegemann, M. Pohl, and K. Turowski, "Challenging big data engineering: positioning of current and future development," presented at the 4th International Conference on Internet of Things, Big Data and Security (IoTBDS), Heraklion, Greece, May. 2019, pp 351-358.

[13] H.H. Altarturi, K. Ng, M.I.H. Ninggal, A.S.A. Nazri, and A.A.A. Ghani, "A requirement engineering model for big data software," presented at the 2nd International Conference on Big Data Analysis (ICBDA), Beijing, China, Mar. 2017, pp 111-117.

[14] NH. Madhavji, A. Miranskyy, and K. Kontogiannis, "Big picture of big data software Engineering: with example research challenges," presented at the IEEE/ACM 1st International Workshop on Big Data Software Engineering (BIGDSE), Florence, Italy, May 2015, pp 11-14.

[15] D. Arruda, N.H. Madhavji, and I. Noorwali, "A validation study of a requirements engineering artefact model for big data software development projects," presented at the 14th International Conference on Software Technologies (ICSOFT), Prague, Czech Republic, Jul. 2019, pp 106-116.

[16] B. Jan, et al., "Deep learning in big data analytics: a comparative study," Journal of Computer Electrical Engineering, 75(1): 275-287, 2019.

[17] A. Haldorai, A. Ramum, and C. Chow, "Editorial: big data innovation for sustainable cognitive computing," Mobile Netw Application Journal, 24(1): 221-226, 2019.

[18] R.H. Hariri, E.M. Fredericks, and K.M. Bowers, “Uncertainty in big data analytics: survey, opportunities, and challenges,” Journal of Big Data, 6(1), 2019.

[19] M. Ghasemi, “What requirements engineering can learn from process mining,” presented at the 1st International Workshop on Learning from other Disciplines for Requirements Engineering (D4RE), Banff, Canada, Aug. 2018, pp 8-11.

[20] K. Lyko, M. Nitzschke, and AC. Ngonga Ngomo, “Big data acquisition,” in New Horizons for a Data-Driven Economy, Springer, Cham, 2016, pp 35-61.

[21] Z. Liu, P. Yang, and L. Zhang, “A sketch of big data technologies,” presented at the 7th International Conference on Internet Computing for Engineering and Science (ICICSE), Shanghai, China, Sep. 2013, pp 26-29.

[22] X. Han, X. Wang, and H. Fan, “Requirements analysis and application research of big data in power network dispatching and planning,” presented at the 3rd Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing Shi, China, Oct. 2017, pp 663-668.

SYSTEMATIC ATTACK SURFACE REDUCTION FOR DEPLOYED SENTIMENT ANALYSIS MODELS

Josh Kalin[1], David Noever[2], and Gerry Dozier[1]

[1]Department of Computer Science and Software Engineering, Auburn University, Auburn, AL, USA
[2]PeopleTec, Inc, Huntsville, AL, USA

ABSTRACT

This work proposes a structured approach to baselining a model, identifying attack vectors, and securing the machine learning models after deployment. This method for securing each model post deployment is called the BAD (Build, Attack, and Defend) Architecture. Two implementations of the BAD architecture are evaluated to quantify the adversarial life cycle for a black box Sentiment Analysis system. As a challenging diagnostic, the Jigsaw Toxic Bias dataset is selected as the baseline in our performance tool. Each implementation of the architecture will build a baseline performance report, attack a common weakness, and defend the incoming attack. As an important note: each attack surface demonstrated in this work is detectable and preventable. The goal is to demonstrate a viable methodology for securing a machine learning model in a production setting.

KEYWORDS

Machine Learning, Sentiment Analysis, Adversarial Attacks, Substitution Attacks.

1. INTRODUCTION

This paper is structured into six separate sections: Introduction, Background, Approach, Evaluation, Future Contributions, and Conclusions.

Sentiment Analysis (SA) [1] is the task of analyzing text to provide a classification such as positive, negative, or neutral for a given sample. SA is subdivided further into categories such as polarity, subject, and toxicity. Companies and organizations use these technologies to moderate their websites, apps, and comment sections [2]. Adversarial attacks, in the context of this work, refer to any input that allows an adversarial actor to trick a classification system. Modern SA Systems use machine learning (ML) and are susceptible to adversarial attacks [3]. Recently, the Natural Language Processing (NLP) community has explored how to create models that can handle bias in training data; for instance, content-aware models are an example of a system that can interpret bias in the data and correctly classify sentiment [4]. Given the challenging nature of SA with this type of data, the goal is to demonstrate a simple and repeatable process for creating a model baseline, attacking the model, and defending against the incoming attacks.

Toxicity Classification [5] is a SA technique to understand the malicious intent of text based on words and content in the message. These SA techniques use ML and Deep Learning (DL) to

David C. Wyld et al. (Eds): SPTM, IPPR, AIS, CSIT, DaMi, AMLA - 2020
pp. 89-99, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100609

classify the toxicity or polarity of a tweet [6]. The first SA technique used in this paper is the Sentiment140 SA API [7]. Sentiment140 originated as a paper from the early 2010s and was later developed into an API by a Stanford Team [8]. Perspective, the second SA API used, is built and maintained by Google's Jigsaw team [9]. The Perspective API is a black box ML model that relies on a transformer and other deep learning technologies to classify sentiment. The Perspective API focuses on toxicity analysis for social media-based comments [10]. Each SA API uses machine learning to provide sentiment classification. We have no connection or insight into the underlying models other than published papers or websites. Further, there are limitations on the number of queries per second and per day.

1.1. Challenges

The dataset used in this work creates a unique challenge. The Conversation AI Team, funded in conjunction with Jigsaw and Google, created a dataset around toxicity, biases, and threats in comment sections [11]. The JigSaw Toxic Bias dataset is a set of publicly released comments augmented with new labels for ML tasks. It has a wide range of different toxicity classifications such as severe toxicity, obscene, identity attack, insult, and threat. In the last year, the Conversation AI team has augmented JigSaw with additional categories including gender, sexual orientation, and religious identity [9]. The new evaluation categories were added to combat inherent biases that are included in the data but do not represent a negative sentiment. This dataset is part of a challenge on the Kaggle competition page for creating the best classification models around toxicity. The top-scoring models used ensembles of DL models to get the highest classification scores [12]. Since the newest classification techniques for this dataset use ML, they are susceptible to adversarial methods [13]. Adversarial Methods demonstrated in the Evaluation section are focused on discovering attacks that negatively affect the classification capability of the underlying system. This paper will focus on the challenge of evaluating the attack surface of a single attack vector and defending the model from this incoming attack.

1.2. Contributions

Single Character attacks vectors are a direct analog to single-pixel attacks in the image domain - for instance, single-pixel attacks have demonstrated effects on classification, reinforcement learning, and other state-of-the-art image technologies [14-15]. We demonstrate the efficacy of single character attacks (1 or many) on these sentiment text classifiers and how to protect the underlying system. These simple attacks can reduce the ability of systems to filter and curate online media platforms. This paper focuses on demonstrating this new architecture to build a baseline of the performance of each API, attack the models with single character substitution/insertion attacks in the text domain, and provide a defense plan for these attacks. The remainder of this paper is as follows: Section 2 presents the related work, Section 3 develops the approach, Section 4 discusses solutions, and Section 5 closes with Conclusions and Future Work.

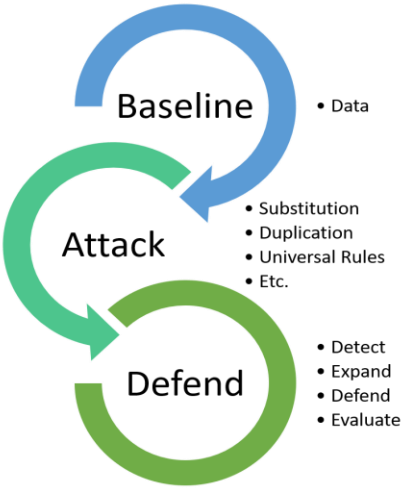


Figure 1: BAD System: Baseline, Attack, Defend for Protecting Machine Learning Models

2. RELATED WORK

There are three key background areas: Build, Defend, Attack Systems from the Cyber Security Domain, the Sentiment Analysis, and black-box models.

2.1. Build, Defend, Attack Systems

Build, Attack, Defend is a construct built around breaking down Cyber Security problems into actionable problem spaces for teams to digest [16]. There are relevant team constructs around Color teams for attacking, defending, and designing systems:

- Red Team: Ethical hacking of a target system
- Blue Team: A group of people building defenses against the attacks on the model
- Green Team: A blend of the red and blue teams where the group will simultaneously run both attacks and create defenses for those attacks [17].

The focus is on applying the green team construct to improve ML model development. By utilizing adversarial design knowledge (Red Team) and model building knowledge (Blue Team), the Green Team can propose defenses to the underlying model designs or production pipelines that secure the model from outside attacks.

2.2. Sentiment Analysis

The field of SA has progressed rapidly due to the expansion of social media platforms. Among internet moderation applications, SA is a required piece of policing social media and comment sections due to the large volume of comments on these sites [18]. With state-of-the-art SA systems improving, it became evident that there is an unintended bias built into the dataset of comments and tweets [10]. Overfitting to words in bias dataset has led to embarrassing results for production SA. The famous example with Perspective API is "I am a gay black woman" which carried a 95% toxicity in 2017 and still contains a toxicity score of 44% as of writing this paper [19]. State-of-the-art SA papers are dominated by ensembles of transformer technologies for this particular problem set [20]. As a note: Each of the SA classifiers inherits weaknesses of the underlying language models used for the classification tasks.

2.3. Black Box Models

A black box model in ML is any model that an end-user only has access to inputs and outputs [21]. Each ML black box model in this work allows an end-user to interact with it through JSON inputs and outputs. There are also limitations to the number of queries per user per system. Two black box SA systems were selected in this paper: Sentiment140 and Perspective API. Sentiment140 is based on a technical report which collected 1.6million tweets to survey ML techniques in the SA domain in 2008 [8]. The Sentiment140 team has maintained the API as a historical benchmark for future SA systems but provides no explicit details on the exact implementation of the API. The Sentiment140 team implemented the paper. In contrast, the Perspective team provides toxicity scores for multiple categories through their API. With Perspective, an end-user can request classification probability scores for each class and can therefore evaluate the efficacy of each attack. The Perspective Team does not provide details on their machine learning models.

3. APPROACH

There are numerous areas of modeling where an adversarial actor can attack. For simplicity, the focus is on inference-based attacks. Attacks on the inference pipeline exploit weaknesses of data used for training and learned weights of the model. For example, there are simple attacks like substitution, replication, and insertion that easily fool current classification models. A recent paper proved universal rules for fooling text-based classification systems are effective for multiple tasks in NLP [22]. The Evaluation section demonstrates a BAD architecture focused on an inference-based attack for each API. Figure 1 shows the general flow of implementing the BAD architecture for an inference attack surface of a SA Model. The following sections discuss each of the Build, Attack, and Defend core components in detail.

3.1. Holistic Approach: Introducing BAD Architecture

Every machine learning team wants to understand the model's vulnerability to adversarial attacks. The BAD Architecture proposes three key steps. First, a team needs to understand the baseline performance of the model by asking questions like the following:

- How does the model act with regular and irregular data?
- Are there known weaknesses or limitations?
- Are those limitations and risks mitigated?

Next, a team needs to understand the impact of each attack by exercising each vulnerability in the model. Last, after understanding the baseline performance and attacking their model, the team will need to propose and implement those defenses to protect their production process. In practice, this entire architecture is repeatable and expandable depending on the scope of the team.

3.2. BUILD a baseline of our target system

A core componentof a ML production system is to understand performance under normal conditions. With the BAD Architecture, each team should also note the known limitations of the model. For instance, some systems do not inherently return real scores for words not in the original training data (example: Word2Vec) [23]. A team must be upfront and understand the impact of design decisions on how a ML system has been designed. To baseline a ML system, it is also important to experiment with data that the system is expected to operate on regularly. If possible, it is also expected to document any edge cases that would be hard for the system to classify. Using an SA black-box model with the JigSaw dataset is a perfect example baseline case for the Build, Attack, Defend Architecture. The data contains toxic edge cases where it is hard to judge the intent of the underlying message. The advantages of creating a systematic baseline are shown in the Evaluation section when edge cases are exploited.

3.3. ATTACK System weaknesses and inefficiencies

Each API has a public page and allows anyone to sign up for basic services. Even with basic access, it is possible to circumvent these systems with a limited number of queries and the Python programming language [24]. Adversarial Character Attacks in the NLP field revolve around changing one or more characters while maintaining the original intent to a human annotator. An Adversarial Attacks using character attacks attempt to direct the decision boundary of the underlying detector in a way that is beneficial to the attacker [25]. For a SA system, this would use substitution attacks to avoid the detection of negative or toxic comments in a social media environment. If bad actors understood how to substitute common character and reduce their toxicity, then it becomes easy for them to use hate speech (as an example). There are two areas in the character attack space applied here: substitution and duplication. The example Attack system demonstrates simple substitution attacks:

1) Create a dictionary that contains vowel to alpha-numeric (for instance e:3)
2) For every vowel in the sentence, replace a single instance from the string and store in an array
3) For every string in the array, evaluate sentiment through public-facing API
4) Evaluate the number of times a single character changed the score or decision made by the black-box model

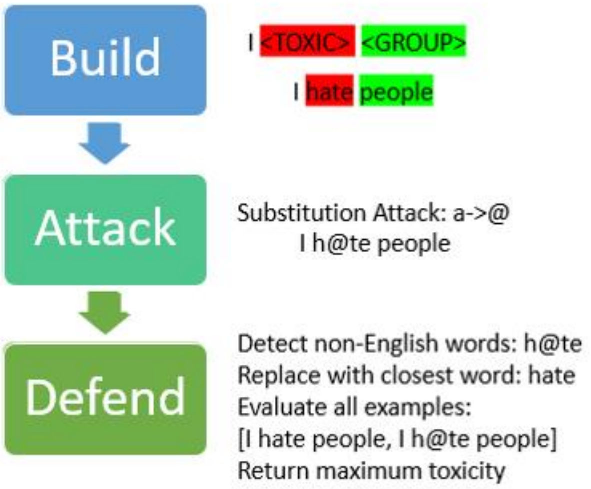
And, for the duplication Attack, the same process is replicated with only a change in the attack vector:

1) Create a dictionary that contains vowels to duplicated vowels (for instance e:ee)
2) For every vowel in the sentence, replace a single instance from the string and store in an array
3) For every string in the array, evaluate sentiment through public-facing API
4) Evaluate the number of times a single character changed the score or decision made by the black-box model

Given the nature of black-box models, each API only provides the probability of toxicity or polarity without additional feedback. With Sentiment140, we are provided three states of polarity: negative, neutral, and positive. There are no percentages of each classification; rather the API simply provides the highest binary classification value. It is only possible to show if a classification can move from one category to another. With Perspective, the actual probability of each classification category is available for each request. Therefore, it is possible to see the decrease or increase in confidence for a given input. The Evaluation section shows the baseline and delta results for each of the attacks.

3.4. Hello World of the BAD Architecture

Figure 2 shows how the system will operate on the simplest incoming toxic phrase. In this example, the 'I hate people' example demonstrates the Build, Attack, and Defend pipeline. Applying the Perspective API, this string scores an 82% toxicity and negative score on the Sentiment140 system. When a Red Team attacks the model with a single character substitution attack of "a" to "@", the toxicity of the comment goes down to 30%. The Green Team's goal is to break apart each attack vector and create a more robust system against adversarial attacks. The Defend section will cover possible strategies for combating simple attacks.


Build
I <TOXIC> <GROUP>
I hate people
Attack
Substitution Attack: a->@
I h@te people
Defend
Detect non-English words: h@te
Replace with closest word: hate
Evaluate all examples:
[I hate people, I h@te people]
Return maximum toxicity

3.5. DEFEND System from Targeted Attack

The Green Team will summarize the baseline and adversarial results to create a defense plan. Typically, there are simple ways to mitigate attacks. For example, a back-end developer can create limitations around what types of requests can be made. Take the commentary systems on a forum: if they use the Perspective API, it would be straightforward to add a few rules to reduce the ability of an attacker to use substitution attacks (Evaluation section covers a basic implementation). When production systems are used as 'download and deploy', creates a wide variety of pointed and effective attacks vectors for adversarial actors. In practice, there are the following crucial steps:

1) Detect: Detect and catch the adversarial text
2) Expand: Expanding the text to include the possible meanings of the originator
3) Defend: Process each result and store for future analysis
4) Evaluate: Check each result and return tune how the team wants the system to respond to attacks

This process cannot stay stationary. Bad Actors are constantly working to find new and inventive ways to break ML systems. The goal of this process is to create development architecture that can be deployed ML model development. In our Evaluation section, we focus on SA and the way we use this system to evaluate the Sentiment140 and Perspective API systems.

4. EVALUATION

Each API provides the ability to send one query per second (1 QPS). There are limitations to the number of adversarial examples we could present to the Perspective API for instance which had a limit of 100 one-second queries. In this instance, a local model is trained and a model is attacked. Then, the potent attacks that fooled the local SA system are used against the black-box model. In practice, adversarial examples were drawn from the training set as a sample of the one hundred top toxic examples for each category of JigSaw. There is a section for Sentiment140 and Perspective where the Build, Attack, Defend Architecture is explained in detail.

4.1. Sentiment140

The Sentiment140 API has a large limit to queries (approximately 5000 items per query per second). There is a maximum limit of around 800,000 scored queries in a given time period (experimentally derived). Experiments are limited to a few permutations of substitution and insertion attacks per input row. In the Build section, the baseline results of the Sentiment140 model with the Jigsaw dataset are covered. In the Attack section, experimental results with simply applying substitution and insertion attacks are explored. The baseline experiments for the JigSaw dataset will be paralleled between the two systems - choose one hundred samples from each toxic category with a 50% or above toxicity and then attack the classification of each toxic row. For Sentiment140, we are given binary results for each experiment. Every returned row provides a category of positive, negative, or neutral. The baseline results are seen in Figure 3 in the Build Section. The classification binary score for these toxic comments is the majority in the neutral and positive categories. For example, 69 percent of the threat category is classified as either neutral or positive. As a note, each toxic input has been annotated by a human to include the label. Sentiment140 does still miss out on a large chunk of the proper classifications of negative for each one of these input rows. The next experiment will show how substitution attacks push the decision boundaries for this model in a different direction.

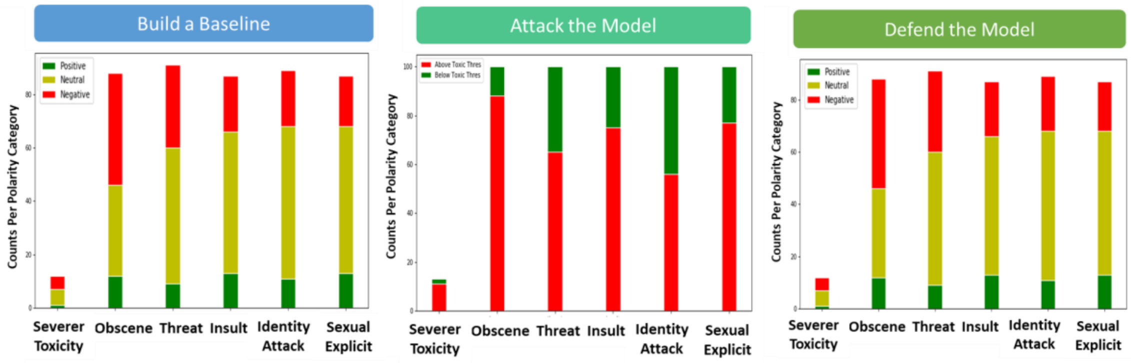


Figure 3: Sentimet140 Performance Summary using the Build, Attack, Defend Development Architecture

4.1.1 ATTACK: Results with Substitution Attacks

The substitution attacks had an interesting effect. The neutral classifications almost universally transformed into negative sentiment. The threat category, for instance, went from 31 percent negative to a 69 percent negative for the one hundred samples included here. In Figure 3, the massive change in the polarity is evident in each category. If the goal was to simply push all polarities to positive, then additional experiments with additional techniques would need to be explored. This work, demonstrates two things:

1) The Sentiment140 algorithm, with this open API, is not well equipped to deal with the bias inherent in modern social media commentary.
2) The decision boundary between neutral and negative is much closer than anticipated with simple substitutions changing neutral polarity into negative polarity.

Sentiment140 was never meant to work with social commentary with this level of toxic bias. Since these are black-box models, there is no opportunity to improve the performance of the underlying system. This highlights the core advantage of applying this architecture. With simple access, the underlying ML model can be evaluated and tested. For defense, both APIs are covered under the Adversarial Attack Surface Reduction section.

4.2. Perspective API

The Perspective API allows one text field per query per second. There is a limit to the number of daily queries but it was not a problem in the experiments. The first step is to create a baseline performance on 100 examples from each of the toxicity categories. Then, attack the same sample of 100 with character attacks (alpha-numeric and duplication) in each category to understand how much degradation can be introduced with simple character attacks.

4.2.1. BUILD: Baseline JigSaw Performance with Perspective

There are two separate experiments run during the baseline stage. A baseline of production models against the Jigsaw data is evaluated. This data is pulled as a sample, straight from the training set, with each toxic category measuring at 50% toxicity or greater for that category (same process as with Sentiment140). If the model were able to classify them appropriately, every one of the examples we pulled would be toxic (red). Our results, shown in Figure 4, show that there is still work to be done in terms of getting full coverage of even just the hundred selected training examples.

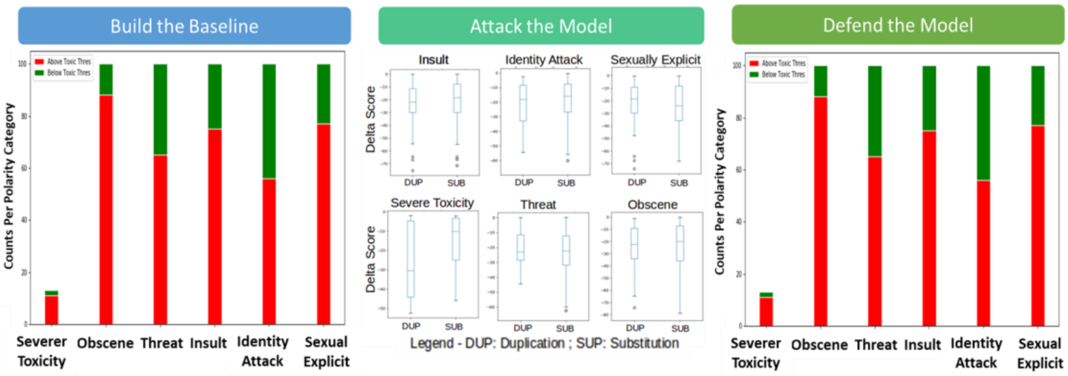


Figure 4: Perspective API Performance Summary using the Build, Attack, Defend Development Architecture

4.2.2. ATTACK: Results with Substitution Attacks

First, for every vowel, substitution and duplication attacks at each vowel position are applied. The Perspective API is robust to these types of simpler attacks, as the delta between the original score and the attacked score was less than 5% in the experiments. Only a single example went below 4% delta in its delta score. This portion of the attack space will need more exploration. One can think about the single character changed once in a string as a single-pixel attack. Only a single pixel is being changed throughout the entire picture. There is another approach to the single character problem - in this attack, consider changing the same character and replace or duplicate it at every instance. The analog to this attack would be changing a certain color pixel throughout the entire image. This exercise is left to another time. The final experiment, highlighted in Figure 4, explores changing vowels to an alpha-numeric or simply duplicating them. Substitution and Duplication attacks are discussed throughout the literature as rudimentary but effective tools when surveying the adversarial surface of these models [22]. In this experiment, the focus was on vowel substitution and duplication attacks on the 100 samples on a per-category basis. In every case, there was at least one example of a -70% reduction in the toxicity score, effectively taking the sentiment from toxic to non-toxic. By using the Attack Surface Reduction steps for our Defend step, it is possible to completely negate the original attacks demonstrated in the last two sections.

4.3. ATTACK SURFACE REDUCTION FOR BOTH APIS

For each of the substitution attacks, there are simple code changes offered for filtering each of these results. In fact, by utilizing these filtering techniques, it is possible to restore the original classification accuracy of the system. Unfortunately, this work does not focus on improving the black box models. The goal is to demonstrate vulnerabilities inherent in these systems and propose an architecture for production systems to protect the efficiency of their systems.

4.3.1. Attack Surface Reduction: Substitution Attacks

This work features two specific types of character attacks - replacement and duplication. The focus is limited to English in this effort although these methods should translate to other languages. First, for detection, the user can detect all non-English words. Multiple models will provide the nearest word or words in a corpus of available words. A simple Defend preprocessing pipeline before inference would be as follows:

1) Find all non-English words by using standard NLP libraries such as NLTK [26]
2) For each non-English word, find nearest neighbor words using algorithms such as Word2Vec [27]
3) Create an array of text with the non-English words replaced with the top N candidates
4) Evaluate the array of text and take the max or min score for classification

In practice, there are commonly misspelled words that can be safely ignored. Using this Defend pipeline can increase the number of candidates for inference but provides robustness to the attacks shown in this paper. As another method, the systems can maintain a set of common substitutions such as alpha-numeric substitutions using alpha-numeric characters or other simple dictionary lookups. Each detected "Attack" should be stored and evaluated by the Green Team periodically to ensure that the pipeline is working as designed.

4.4. Limitations

In each example, the attacks were simple character to character mappings. Zero-Day attacks in the cyber realm refer to attacks that are not yet protected against and allow a hacker unfettered access to a system. In the machine learning realm, there are adversarial ‘zero-day’ attacks that are manipulating the output of the model. These Zero-Day attacks are difficult to anticipate and protect against in practice. This architecture currently relies on known attacks on models for protections and does not actively search an adversarial surface for the model.

5. Conclusions and Future Work

This work demonstrates that deployed sentiment models are susceptible to simple substitution attacks on single characters and can be effectively defended from each substitution attack using the BAD architecture. Because these substitutions are simple character to character mappings, they are mitigated by detecting non-English words, creating candidates for sentiment analysis, and taking the maximum toxicity in our examples. Further work in this area will focus on looking at model attacks like weight poisoning attacks on classification systems.

Weight Poisoning Attacks on Pre-trained Models [28] is a recent paper that uses vulnerabilities in pre-trained models and strikes me as dangerous to all black box models that are not actively defending against those types of tasks. A future direction could be to develop a data augmentation method or model structure that makes weight poisoning attacks reduces the efficacy of weight poisoning attacks, During the defend phase, automated methods for detecting and correcting poisoned words could use transformer models to find and propose corrected word.

Acknowledgments

Special thanks to the medical folks out there during the COVID-19 pandemic.

References

[1] R. Feldman, “Techniques and applications for sentiment analysis.” Commun. ACM, vol. 56, no. 4, pp. 82–89, 2013.

[2] A. Kumar, T. M. Sebastian et al., “Sentiment analysis: A perspective on its past, present and future,” International Journal of Intelligent Systems and Applications, vol. 4, no. 10, pp. 1–14, 2012.

[3] J. Gao, J. Lanchantin, M. L. Soffa, and Y. Qi, “Black-box generation of adversarial text sequences to evade deep learning classifiers,” in 2018 IEEE Security and Privacy Workshops (SPW). IEEE, 2018, pp. 50–56.

[4] L. Gao and R. Huang, “Detecting online hate speech using context aware models,” arXiv preprint arXiv:1710.07395, 2017.

[5] J. Guberman, C. Schmitz, and L. Hemphill, “Quantifying toxicity and verbal violence on twitter,” in Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion. ACM, 2016, pp. 277–280.

[6] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” arXiv preprint arXiv:1810.04805, 2018.

[7] W. Medhat, A. Hassan, and H. Korashy, “Sentiment analysis algorithms and applications: A survey,” Ain Shams engineering journal, vol. 5, no. 4, pp. 1093–1113, 2014.

[8] A. Go, R. Bhayani, and L. Huang, “Twitter sentiment classification using distant supervision,” CS224N Project Report, Stanford, vol. 1, no. 12, p. 2009, 2009.

[9] P. Team. (2019) ”Jigsaw toxic bias FAQ”. [Online]. Available: https://www.kaggle.com/c/jigsaw-unintended-bias-intoxicity-classification/overview/faq

[10] “Perspective API,” https://www.perspectiveapi.com//home, accessed: 2019-11-11.

[11] D. Borkan, L. Dixon, J. Sorensen, N. Thain, and L. Vasserman, “Nuanced metrics for measuring unintended bias with real data for text classification,” in Companion Proceedings of The 2019 World Wide Web Conference. ACM, 2019, pp. 491–500.

[12] C. AI, “Jigsaw unintended bias in toxicity classification,”https://www.kaggle.com/c/jigsaw-unintended-bias-in-toxicity classification/overview, 8 2019.

[13] L. Huang, A. D. Joseph, B. Nelson, B. I. Rubinstein, and J. D. Tygar, “Adversarial machine learning,” in Proceedings of the 4th ACM workshop on Security and artificial intelligence. ACM, 2011, pp. 43–58.

[14] J. Su, D. V. Vargas, and K. Sakurai, “One pixel attack for fooling deep neural networks,” IEEE Transactions on Evolutionary Computation, 2019.

[15] T. Chen, J. Liu, Y. Xiang, W. Niu, E. Tong, and Z. Han, “Adversarial attack and defense in reinforcement learning-from ai security view,” Cybersecurity, vol. 2, no. 1, p. 11, 2019.

[16] M. O. Leary, Cyber Operations: Building, Defending, and Attacking Modern Computer Networks. New York, NY: Appress, 2015.

[17] A. C. Wright, “Orange is the new purple,” 2017, blackHat USA 2017. [Online]. Available: https://www.blackhat.com/docs/us-17/wednesday/us-17-Wright-Orange-Is-The-New-Purple-wp.pdf

[18] R. Pandarachalil, S. Sendhilkumar, and G. Mahalakshmi, “Twitter sentiment analysis for large-scale data: an unsupervised approach,” Cognitive Computation, vol. 7, no. 2, pp. 254–262, 2015.

[19] J. West, “I tested 14 sentences for ”perceived toxicity” using Perspectives.” 2017. [Online]. Available: https://twitter.com/jessamyn/status/900867154412699649?lang=en

[20] JigSaw. (2019) ”Perspective”. [Online]. Available: https://www.perspectiveapi.com

[21] L. Ljung, “Black-box models from input-output measurements,” in IMTC 2001. Proceedings of the 18th IEEE instrumentation and measurement technology conference. Rediscovering measurement in the age of informatics (Cat. No. 01CH 37188), vol. 1. IEEE, 2001, pp. 138–146.

[22] D. Li, D. V. Vargas, and S. Kouichi, “Universal rules for fooling deep neural networks based text classification,” arXiv preprint arXiv:1901.07132, 2019.

[23] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,” in Advances in neural information processing systems, 2013, pp. 3111–3119.

[24] G. Van Rossum and F. L. Drake, The python language reference manual. Network Theory Ltd., 2011.

[25] J. Ebrahimi, A. Rao, D. Lowd, and D. Dou, “Hotflip: White-box adversarial examples for text classification,” arXiv preprint arXiv:1712.06751, 2017.

[26] E. Loper and S. Bird, “NLTK: the natural language toolkit,” arXiv preprint cs/0205028, 2002.

[27] P. Laskov and M. Kloft, “A framework for quantitative security analysis of machine learning,” in Proceedings of the 2nd ACM workshop on Security and artificial intelligence. ACM, 2009, pp. 1–4.

[28] Kurita, Keita, Paul Michel, and Graham Neubig. "Weight Poisoning Attacks on Pre-trained Models." arXiv preprint arXiv:2004.06660 (2020).

AUTHORS

Josh Kalin is a physicist and data scientist focused on the intersections of robotics, data science, and machine learning. Josh holds degrees in Physics, Mechanical Engineering, and Computer Science.

David Noever has research experience with NASA and the Department of Defense in machine learning and data mining. He received his Ph.D. from Oxford University, as a Rhodes Scholar, in theoretical physics.

Gerry Vernon Dozier, Ph.D., is the Charles D. McCrary Eminent Chair Professor in the Department of Computer Science & Software Engineering at Auburn University. Dr. Dozier is the director of the Artificial Intelligence & Identity Research (AI2R) Lab. He is currently applying his research expertise in the areas of Artificial Intelligence, Machine Learning, Behavioral Analytics, and Evolutionary Computation to the areas of Cyber Identity Protection and Cyber Security. Dr. Dozier has published over 140 conference and journal publications. Dr. Dozier earned his Ph.D. in Computer Science from North Carolina State University.

AUTOMATION OF PURCHASE ORDER IN MICROSOFT DYNAMICS 365 BY DEPLOYING SELENIUM

Vijay Biju and Shahid Ali

Department of Information Technology, AGI Institute,
Auckland, New Zealand

ABSTRACT

Regression testing is very important for dynamic verification. It helps to simulate a suite of test cases periodically and after major changes in the design or its environment, to check that no new bugs were introduced. Evidences regarding benefit of implementing automation testing which includes saves of time and cost as it can re-run test scripts again and again and hence is much quicker than manual testing, providing more confidence in the quality of the product and increasing the ability to meet schedules and significantly reducing the effort that automation requires from testers are provided on the basis of survey of 115 software professionals. In addition to this, automated regression suite has an ability to explore the whole software every day without requiring much of manual effort. Also, bug identification is easier after the incorrect changes have been made. Genius is going through continuous development and requires testing again and again to check if new feature implementation have affected the existing functionality. In addition to this, Erudite is facing issue in validation of the Genius installation at client site since it requires availability of testers to check the critical functionality of the software manually. Erudite wants to create an automated regression suite for Genius which can be executed at client site for checking the functionality of the software. In addition to this, this suite will also help the testing team to validate if the new features which have been added to the existing software are affecting the existing system or not. Visual studio, Selenium Webdriver, Visual SVN and Trello are the tools which have been used to achieve the creation of automation regression suite. The current research will provide guidelines to the future researchers on how to create an automated regression suite for any web application using open source tools.

KEYWORDS

Automation testing, Regression testing, Visual Studio, C#, Selenium Webdriver, Agile- Scrum

1. INTRODUCTION

Independent Advisory Services Limited (IAS Ltd) is a software company which was founded in 2003 which aims to be the most successful, creative and path breaking consulting agency in Australia and New Zealand. The aim of the company is to approach the clients with unique ideas to develop customized and different strategies. The main strategies offered by the company are Enterprise resource planning (ERP), digital strategies and operational business strategy. Demo data is the basic data set which is released for implementation of support and demonstration purposes. The current demo data set supports verticals like Retail, Distribution, Service Industries, Public Sector and Discrete & Process Manufacturing. Demo data supports around 40 languages around 16 countries. In Dynamics 365 Field Service, a purchase order (P.O.) is created to purchase the products to sell to a customer in a work order. Company must automate the

David C. Wyld et al. (Eds): SPTM, IPPR, AIS, CSIT, DaMi, AMLA - 2020
pp. 101-114, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100610

purchase order creation for a liquor vendor in Microsoft Dynamics 365 and develop the test cases against the requirement. For automating the process, we must select a tool which is an open source, stable and supports all the programming languages. While studying the New Zealand market we found that Selenium will be the ideal tool for automating the purchase order creation of the various features to use like plugins, reporting etc. After analysing the requirements, we have developed a hybrid framework where all types of concept like test driven framework, Keyword driven framework etc all comes together which is easy to use and can be used for the future sprints. We have used two kinds of reporting which is TestNG, Extent reports for sharing the project execution results to the Stakeholders. For attaining this goal company must adopt an agile methodology where all the team members and individuals effectively prioritize the features and works to attain the product quality in the entire development phase.

IAS Ltd is to automate the purchase order creation and make a payment to the vendor. For attaining this goal, we must develop the test cases for automating the purchase order in Dynamics 365. Test results for each test cases will be generated through TestNG and extent reports. This reporting helps to identify the passed and failed scenarios with a simple pie chart showing the progress of the project execution. Each test cases will be written in separate methods and main class will call the methods while execution. A hybrid framework has been developed where test data is read from an excel file and failed test cases is captured by a screenshot method. Automation framework will be constructed as page object model in Maven which helps to maintain the framework and changes can be easily tackled in this approach.

This research project report is organized as follow: Section 2 focuses on the literature review of various studies concentrating on automation regression. Section 3 is focused on the research methodology for this research project. Section 4 of this research is focused on project execution results. Discussion to results of this project are provided in section 5. Section 6 is dedicated towards the future work recommendations. Finally, in section 7 conclusion to the research is provided.

2. LITERATURE REVIEW

In the past different researches had been conducted for designing of automating framework for websites as discussed below:

A study was conducted on Analysis and Design of Selenium WebDriver by [7]. In this study they use designed and implemented automation testing framework for testing web applications using selenium web driver. In this framework tester can easily write their test cases in less time. The developed framework is helpful for developer to analyse and maintain their code due to screen shot property of framework and helps to generate test report which helps in identifying the passed and failed test cases as well as maintaining the test suite.

Another study was conducted about the behaviour driven Test Automation Framework [19]. This study focuses on how the user layer, technical, business layers and data layer constraints can be implemented and maintained without costing and can support test automation of all the different layers of a three-tier architectural system which will improve the test coverage, quality and reliability of the software system.

A research on focuses on the automation testing tools was proposed by [12] to enhance design and execution activity, challenges faced by non-automation tester in executing automation scripts and executing automation scripts using TestNG and its disadvantages and the proposed web application that nullify the problems faced by manual testers which reduce time on set-up environment to execute test scripts and overcome consequences of execution in TestNG.

Similar research was done by [9] to perform automation testing using "Selenium WebDriver" with this data driven framework it separates data to code for the purpose of reusability.

Another research was done by [10] to repair the test suite when modification is done to the application when test cases are to be changed.

A research that focuses on how Selenium tool will fasten the execution process and reduce the cost of the test execution process by [11]. The paper gives the information about how to focus on frequent integration which is the main feature of agile.

Another research by [12] develop an architecture that can be used repeatedly through the common application of the agile software and test action planning for automating the acceptance test.

A research by [13] that focuses on developing a framework which aims at reducing implementation and maintenance costs of automation tests and focuses on creating tests which can be easily understand by testers or stakeholders without previous coding knowledge. When a test fails while locating a page element which is not there due to outdated test script, Smart Driver automatically analyses the element which is moved to another place in the page.

Another study by [14] focuses on the best testing tool for checking the functionalities, security and performance with minimum cost and the paper fails on how to successfully implement the tools.

A study by [15] focuses on a new automated testing framework for testing the web applications which improves the automating process. The proposed framework saves around three-fourth of the total effort involved in the automation process using old automation and 21% compared to using Selenium IDE.

A research by [16] focuses on exploring various types of software testing, techniques and tools to compare manual testing against automation testing.

A research by [17] that focuses on a model for automated agile testing, and an working framework developed on the testing of a Web application and results are evaluated using the agile testing model, and there is a comparison between waterfall and agile models.

Similar research by [18] that focuses on case study by surveys were used for data collection and challenges connected to the testing process for a complex project environment and unscheduled releases were identified, on the identified results its concluded that the described approach addresses well the described issues and furthermore efficient testing environment that combines a number of test frameworks like JUnit, Selenium with custom-developed simulators is presented.

3. RESEARCH METHODOLOGY

Research methodology for the project has been discussed below.

3.1. Agile Methodology

Agile is the powerful methodology used in the industry today with 52% of companies reporting that more than half of their teams are using agile practices [24]. At the same time, developing secure software is extremely important given the extensive spread of security exploits. NIST reported around 16,000 software vulnerabilities across the industry in 2018 [6]. For the automation of purchase order creation company has adopted the agile methodology. Agile methodology is a method which has continuous iteration of development and testing throughout the software development lifecycle (SDLC) of a project. To work on test automation, a proven test methodology and an automation framework must be used. However, using Hybrid automation framework requires to learn the techniques of the test tool used and time and effort is required for customizing [5]. Automation of purchase order follows agile methodology as agile methodology is easy for implementing. Through this approach, application is built in short cycles, thereby ensuring reliability for timely releases. This results in building, testing and releasing the software faster and more frequently [23]. The scrum methodology has its benefit as it provides increased customer satisfaction with necessary responsiveness for change requests [22].

3.2. Task Schedule

Table 1 represents the schedule of this project in Independent Advisory Services for the automation process for purchase order creation. The whole process was for 5 weeks where each week was labelled as one sprint. In the initial stage of project, we discuss the project and requirements which we have to work on. As part of the first sprint we come across the Microsoft dynamics 365 website where all the procedures were defined clearly. Dynamics website is different from normal web page as it was hard to inspect the elements to find the x path for automation purpose, so the site contains all the shortcuts explained detailly. In sprint 2 we had analysed the New Zealand market we compared and selected the tool which was suitable for automating the purchase order creation. In sprint 3 we had set up the environment by downloading all the jar files need to run the automation scripts. Sample scripts are developed which only contains the needy x paths and check whether its working fine. In the 4th sprint we developed a hybrid framework which is the standard industry model which will be easy to maintain by adding configuration, helper and property file storing all the necessary data and methods to run the automation process. In the final sprint we are writing the report of the automation process for purchase order creation and the power point final presentation.

Table 1: Task Schedule

Sprint	Actions	Start Date	End Date	Duration
Sprint 1	Discussion of the project	October 28	October 30	15 hr
	Analysing the Microsoft dynamics site	October 31	November 2	15 hr
Sprint 2	Exploratory testing on Microsoft dynamics 365 for inspecting the shortcuts available.	November 2	November 5	15 hr
	Review the requirements and selecting the tool for automation	November 6	November 8	15 hr
Sprint 3	Set up the environments	November 9	November 10	15 hr
	Sample scripts of the	November 11	November 13	15 hr

	purchase order creation			
Sprint 4	Developing the Hybrid framework	November 14	November 16	15 hr
	Working on different reporting facility available	November 17	November 21	15 hr
Sprint 5	Report writing	November 22	November 24	15 hr
	Presentation of the report	November 25	November 30	15 hr

3.3. Test Cases

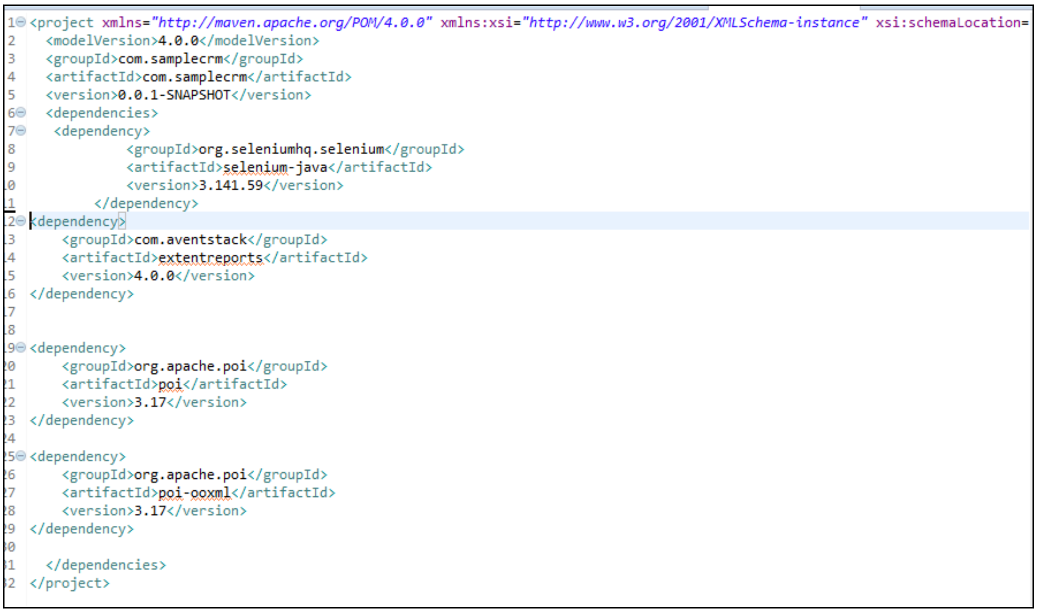
Table 2 shows the test cases of purchase order creation for this project.

Testcase ID	Test Scenario	Pre-Conditions	Test Steps	Expected Results	Actual Results	Test Results
TC_1	Verify login and Password.	Browser Launched and Navigate to https://democontosodatadevaos.sandbox.ax.dynamics.com/?cmp=USMF&mi=DefaultDashboard	Enter the valid Username and Password successfully.	Login must be successful.	Login successfully done.	Pass
TC_2	Create new purchase order	Login Successful	1.Click new order in left most part of Dashboard. 2.Enter the Vendor account and site name. 2.Save the details after entering mandatory fields.	Purchase order must be successful.	Purchase order Creation successfully done.	Pass
TC_3	Search Item number	Login Successful	Enter the associated item number for the vendor	Enter the associated item number for the vendor	Item number must be added successfully	Pass
TC_4	Save and confirm Purchase order	Login Successful	Click the save button and confirm the purchase order.	Click the save button and confirm the purchase order.	Purchase order Creation must be successful.	Pass

Table 2: Test Cases of purchase order creation

3.4. Build Tool

The build tool is used to set up the project that is essential to run the java code for the whole java project. Build tool will generate a source code, compiling code, packaging code to a jar etc. For automating the purchase order creation, we have used the build tool maven. Maven provides a common platform to perform the activities which makes developers task easy. The core of any maven project is the pom.xml, where all the information's are stored. Most of the Integrated Development Environments (IDE) are available for tools like Eclipse, NetBeans, IntelliJ etc. Maven stores all jars like selenium standalone, TestNG, Extent reports and Apache POI for the project. Library jar is placed in central repository from where maven downloads all the dependency jar for the automation process.



```
<project xmlns="http://maven.apache.org/POM/4.0.0" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation=
  <modelVersion>4.0.0</modelVersion>
  <groupId>com.samplecrm</groupId>
  <artifactId>com.samplecrm</artifactId>
  <version>0.0.1-SNAPSHOT</version>
  <dependencies>
   <dependency>
            <groupId>org.seleniumhq.selenium</groupId>
            <artifactId>selenium-java</artifactId>
            <version>3.141.59</version>
        </dependency>
<dependency>
    <groupId>com.aventstack</groupId>
    <artifactId>extentreports</artifactId>
    <version>4.0.0</version>
</dependency>


<dependency>
    <groupId>org.apache.poi</groupId>
    <artifactId>poi</artifactId>
    <version>3.17</version>
</dependency>

<dependency>
    <groupId>org.apache.poi</groupId>
    <artifactId>poi-ooxml</artifactId>
    <version>3.17</version>
</dependency>

  </dependencies>
</project>
```

Figure 1: Pom.xml used in the Project

3.5. Hybrid Framework

Hybrid Framework has the capability to handle many test cases and it can produce accurate results as per the test case. Developed Hybrid framework completely reduces the manual dependency in automation testing. [4]. The main advantage of test automation comes from fast, precise execution of a set of tests after some changes have been made to a web application [21]. The aim of (test data management) TDM is to improve the effectiveness and reduce the time and cost for testing. [2].

Key features in implementing Hybrid Framework with Selenium:

- Stores the input test data in Excel file.
- Can store the environment related information in a property file.
- Store various objects in the applications where user need to access in object repository file.
- Test suite will contain the verifying tasks mentioned in the requirements.
- Executing on different browsers when needed.
- Generated screenshots to capture the passed/failed test cases.

- Multiple Report generation through TestNG, Extent Reports [2].

A framework is a set of assumptions, concepts and practices that needs to be followed. There are many components of framework:

- Test case standardization: Every test must be in a proper format, in order to achieve this, we must follow a design pattern so for the automation of purchase order creation we are using page object model (POM).
- Logs: For each execution we have to generate the logs to check what went wrong so generation of logs are of primary importance.
- Test data and configuration utility: The most important part of the framework is the test data without the test data we cannot fill the mandatory fields while automating, so hard coring the values are not a proper way in Industries because in the further sprints if we want to change the data we have to keep the test data and utility file separately.
- Helper or Utility library: While working on different pages the excel data, file code to read, different browser action will be changed for accessing these libraries for accessing these kinds of information will be stored in Helper and utility libraries.
- Test Execution engine: We must define how our final test will run, which build tools are we using or for accessing libraries whether we are using any continuous integration tools like Jenkins for configuration. For the automation purpose we are using the build tool Maven and the proper way of using like add all the needful dependency in the pom.xml.
- Reporting: The main component of the framework is reporting without proper reports we cannot show the stakeholders the test execution results. In a proper report we have all the details about the methods we made against the test cases, how much time it takes to execute the test cases and a simple diagram showing the pass and fai status.

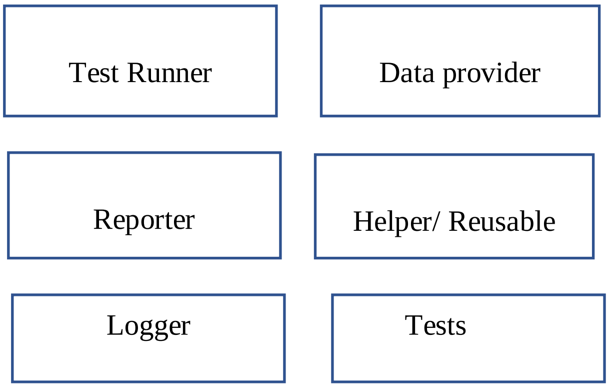


Figure 2: Basic framework design

3.6. Reporting

For the formation of robust Hybrid framework most important element is Reporting. A good test report serves as an overview of the project which helps in finding bugs and send the test report to the Stakeholders to show the nutshell of the test case execution.

Main features of the test reports are:

- Short and precise.
- Show the number of passed and failed test cases.
- Show a pictorial diagram of test execution results.

- Must support a format which can be shared by email or integrate with Continuous Integration (CI) tools like Jenkins/Bamboo..

Show the test coverage of application/module under test.

3.7. Generated Reports in Selenium Web Driver

As an open source Selenium Web driver does not have a built-in reporting feature, as we have added the Maven dependency for Extent Report and TestNG for generating reports.

3.7.1. Testing Reporting

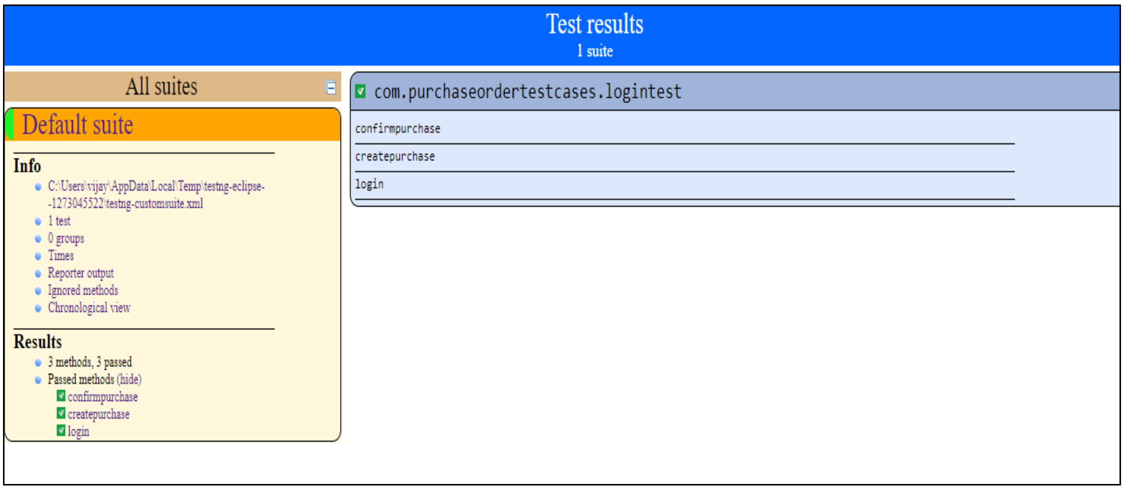
TestNG is a testing framework modified from JUnit and NUnit having additional features which is easy to handle. TestNG is an open source automated testing framework, where NG means Next Generation. TestNG gives the developer to write more flexible and powerful test scripts which borrows from Java Annotations to define tests in a real production environment. Reporting reduces the problems faced by manual testers and time spent on initial set-up activity to carryout test scripts execution and overcome disadvantages of execution using TestNG [1].

TestNG Features

- Supports annotations.
- Supports testing integrated classes.
- Flexible runtime configuration.
- Supports Dependent test methods, load testing, and limited failure.
- Flexible plug-in API.
- Supports multi-threaded testing.

JDK	1.5 or above.
Memory	No special requirements
Disk Space	No special requirements
Operating System	No special requirements

Table 3: System Requirements



Test results
1 suite

All suites

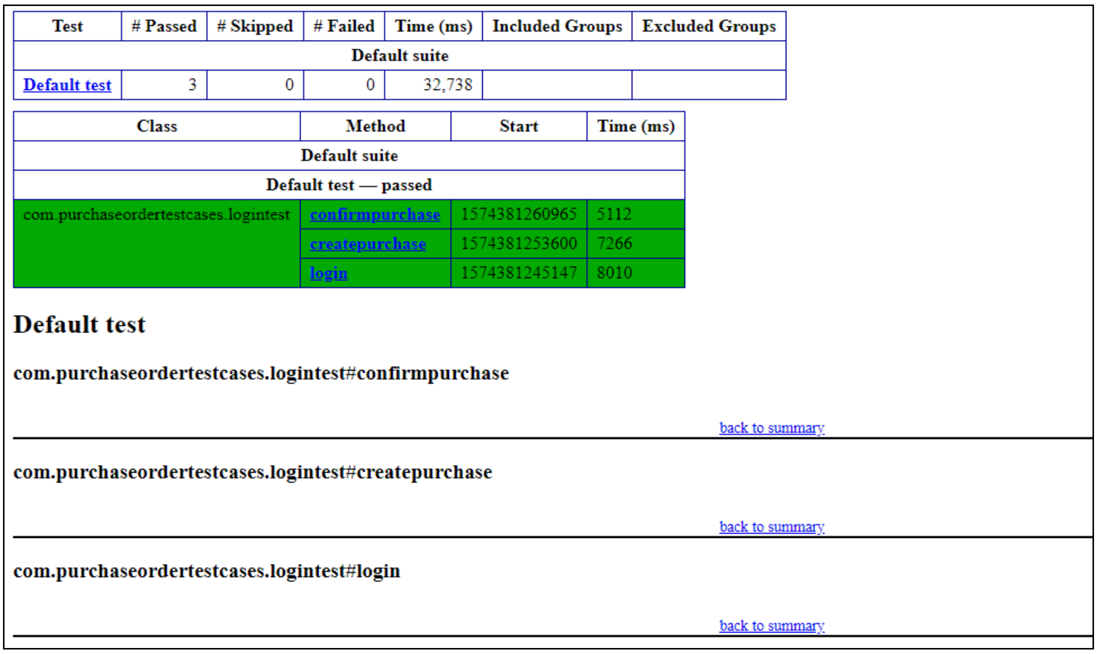
Default suite

Info
- C:\Users\vijay\AppData\Local\Temp\testng-eclipse--1273045522\testng-customsuite.xml
- 1 test
- 0 groups
- Times
- Reporter output
- Ignored methods
- Chronological view

Results
- 3 methods, 3 passed
- Passed methods (hide)
  - ☑ confirmpurchase
  - ☑ createpurchase
  - ☑ login

☑ com.purchaseordertestcases.logintest
confirmpurchase
createpurchase
login

Figure 3: Index.html



Test	# Passed	# Skipped	# Failed	Time (ms)	Included Groups	Excluded Groups
Default suite						
Default test	3	0	0	32,738		

Class	Method	Start	Time (ms)
Default suite			
Default test — passed			
com.purchaseordertestcases.logintest	confirmpurchase	1574381260965	5112
	createpurchase	1574381253600	7266
	login	1574381245147	8010

Default test

com.purchaseordertestcases.logintest#confirmpurchase

back to summary

com.purchaseordertestcases.logintest#createpurchase

back to summary

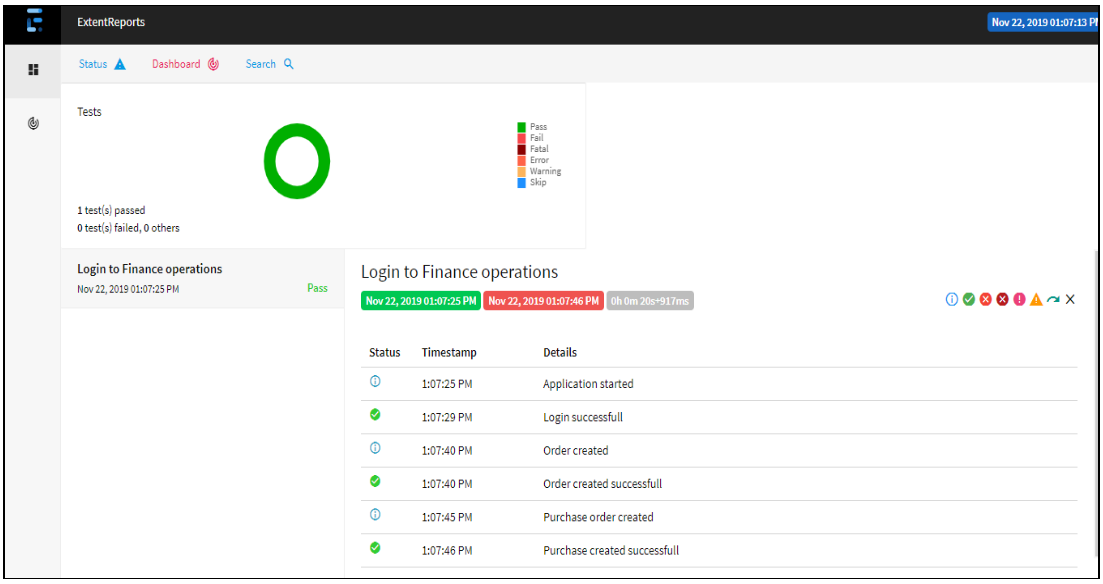
com.purchaseordertestcases.logintest#login

back to summary

Figure 4: Output of TestNG

3.7.2. Extent Reports

Different types of reporting are available in selenium, but for customizable report generation which must be shared with Stakeholders we have used extent reports which can be integrated into Selenium WebDriver using JUnit and TestNG frameworks. Extent Reports have more advantages when compared to the reports generated through JUnit and TestNG such as pie chart representation, test stepwise report, adding screenshots etc at every step and an attractive GUI which show the pass and fail test cases which can be shared to stakeholders.



1 test(s) passed
0 test(s) failed, 0 others

Login to Finance operations
Nov 22, 2019 01:07:25 PM Pass

Login to Finance operations
Nov 22, 2019 01:07:25 PM | Nov 22, 2019 01:07:46 PM | 0h 0m 20s+917ms

Status	Timestamp	Details
ⓘ	1:07:25 PM	Application started
✓	1:07:29 PM	Login successfull
ⓘ	1:07:40 PM	Order created
✓	1:07:40 PM	Order created successfull
ⓘ	1:07:45 PM	Purchase order created
✓	1:07:46 PM	Purchase created successfull

Figure 5: Output report of Extent Report

4. DISCUSSION

While automating the purchase order creation we have faced many problems in this testing life span. The problems and solution found are discussed below:

- In the initial stage of automation testing of the purchase order creation we found that x-path is not detectable in the Microsoft dynamics site so we went to the site of Microsoft dynamics where there is a list how to do the basic keyboard shortcuts details are available. Control + Shift + I is the shortcut key for finding the x-path.
- After writing the scripts and while running it may take a certain time limit to load the dynamics web page. So, to overcome we must put implicit wait for the page to fully load to pick the x path for automation process.
- While running the automation scripts the browser may stuck at some point so while clearing the cache or closing the automated browser, we can solve this problem.
- Google web driver may be updated at every point so we have to download the new version according to the current google version we are using, else we use an if case either to select any browser like Firefox, Internet explorer so in the configuration file if you are changing the browser name other than chrome it will run in the selected browser.

4.1. Comparative Analysis for Automation Tool Selection

In order to attain maximum robust framework which is stable and easy to maintain we have to search for the various tools for automating the application and compared the best two, Selenium and UFT(QTP) and select the best one for automating the purchase order creation.

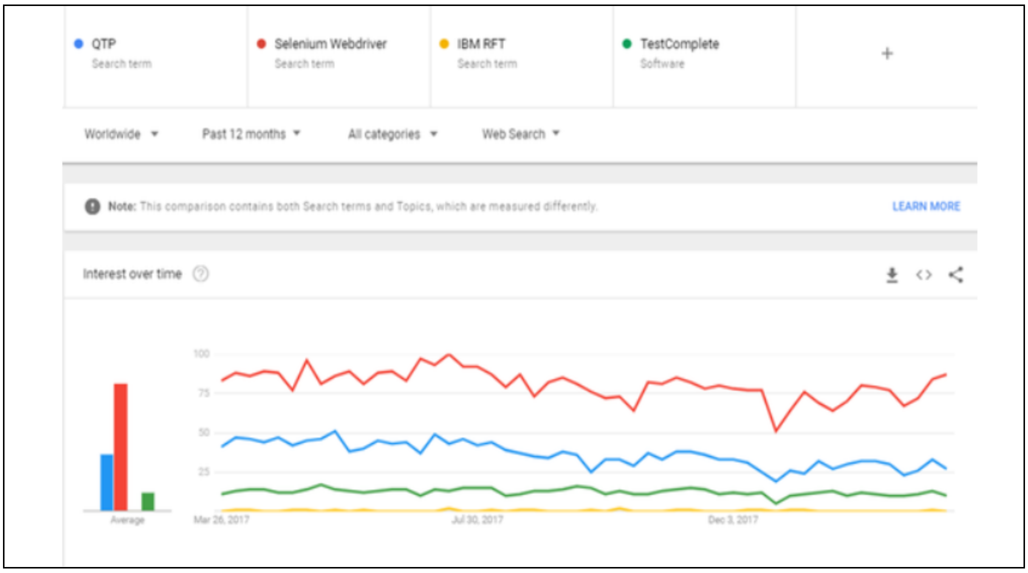
Table 4 represents the comparison of Selenium and Microfocus-UFT.

FEATURES	SELENIUM	UFT(QTP)
Cost	Free tool/open source. No issues in licensing or renewal. Just need to download it and use	Paid tool and will cost an average of $3200, UFT is available as seat-based and concurrent which will be more expensive.
Support	Since it’s an open source no professional support is available.	Since it is a paid tool proper support team is available.
Application Type	Inbuilt, selenium supports only Web Applications. It recognizes the elements on screen using id, CSS selector, xpath.	It supports web, mobile, API, hybrid, RPA, and enterprise application.
Languages supported	Java, C#, Ruby, Python, Perl PHP, JavaScript, R etc.	VBS (Visual Basic Script)
Supported Browsers	IE, Firefox, Chrome, Safari, Opera, Headless browsers.	Chrome, Firefox, Safari, IE, and Edge.
Coding skills	Good knowledge of programming language is needed for each Binding.	Less programming knowledge is required as it offers keyword-driven testing which simplifies test creation and maintenance.

Test Report	Selenium has to download the necessary plugins to generate test reports.	Default test reports are generated.
Performance Testing	Selenium is not used for performance testing but can be integrate with JMeter, to run your selenium scripts for performance testing.	We can mimic the user actions in UFT.
Tools Integration	Can be integrated with paid or free tools	Can be integrated limited paid tools

Table 4: Selenium vs UFT

According to google trends survey globally the mostly used is Selenium. Selenium is a one of the efficient open-source automated testing tool which provide a stable testing framework for testing a wide variety of applications and exporting scripts in almost every language including java, .net, c#. The main feature of Selenium is that it supports different browsers for executing the test cases [20]. In this project Selenium Web driver is used to automate the purchase order and to illustrate the use of selenium tool in combination with other tools like the Maven, TestNG, etc., for more easier approach to testing and to improve the quality of testing process [19]. The inclusion of the Page Object pattern has demonstrated that it will be very effective in end-to-end testing. Page objects are classes which abstracts the web pages into required business functions that can be called by the test cases. By decoupling source code from web page information, test cases made for the web pages are more readable and easier to maintain [21]. The main use of selenium for automation purpose is the flexibility. Selenium features like regrouping and refactoring helps the developers and testers to maintain and quickly change the code.



Worldwide ▾ Past 12 months ▾ All categories ▾ Web Search ▾

Note: This comparison contains both Search terms and Topics, which are measured differently.

Figure 6: Automation Tools usage Globally

Selenium is selected when compared to the other tool because of the following reasons:

- We can automate the purchase order creation at free of cost.
- Developed hybrid framework is easy to maintain for the future sprints if required.
- For scheduling the jobs, we can integrate with Jenkins.
- Various style of Reporting like Extend reporting, TestNG can be done with Selenium.

- Failed test cases are identified through screenshots which has included in the framework.

5. FUTURE RECOMMENDATIONS

Our task about the automation of creation of purchase order in Microsoft Dynamics 365 in selenium web driver. The development process for the automation can be performed by a Behaviour Driven Development (BDD) either by using Cucumber or Spec flow. The acceptance criteria can be added in the feature file so that team members who does not know about the coding like Stakeholder, Business Analyst can understand how the build is made against the requirements. Step definition file is made against each step in the feature file which will be easy to correct the error.

6. CONCLUSION

Automation empowers enterprises to reduce resource allocation, reduce cost to company (CTC), enables business that ensures customer satisfaction. In this execution, execution of automating the purchase order creation for the company Independent Advisory Services has been discussed. Selection of optimum tool from the New Zealand market for automating the purchase order has been discussed. Hybrid Framework has been developed for the automation process and how the framework is used for future maintenance in the upcoming sprints. Different methods have been used for report generation have been discussed. Necessary recommendation and the problems faced during this project has been mentioned. The proposal concluded about what are the testing activities that is to be done for automating the purchase order and to identify test cases and prepare the test scripts and generate the HTML reports.

REFERENCES

[1] Jain, C. R., & Kaluri, R. (2015). Design of automation scripts execution application for selenium webdriver and test NG framework. ARPN J Eng Appl Sci, 10, 2440-2445.

[2] Bajaj, K. S. (2018). Hybrid Test Automation Framework for managing Test Data. International Journal of Pure and Applied Mathematics, 118(9), 265-277.

[3] Bajaj, H. (2015). Choosing the right automation tool and framework is critical to project success. Infosys Limited.

[4] Lamba, S., Rishiwal, V., & Rana, A. (2015). An automated data driven continuous testing framework. International Journal of Advanced Technology in Engineering And Science, 3(1).

[5] Shim, J. A., Kwon, H. J., Jung, H. J., & Hwang, M. S. (2016). Design of acceptance test process with the application of agile development methodology. International Journal of Control and Automation, 9(2), 343-352.

[6] NIST (2019). National Vulnerability Database.

[7] Gojare, S., Joshi, R., & Gaigaware, D. (2015). Analysis and Design of Selenium WebDriver Automation Testing Framework. Procedia Computer Science, 50, 341–346.

[8] Subramanian, R., Chen, N., & Zhu, T. (2017). Behavior Driven Test Automation Framework. In Proceedings of the International Conference on Software Engineering Research and Practice (SERP) (pp. 17-23). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

[9] Chandraprabha, C., Kumar, A., & Saxena, S. (2015). Data Driven Testing Framework using Selenium WebDriver. International Journal of Computer Applications, 118(18), 18–23.

[10] Leotta, M., Clerissi, D., Ricca, F., & Spadaro, C. (2013). Comparing the maintainability of selenium WebDriver test suites employing different locators: a case study. Proceedings of the 2013 International Workshop on Joining AcadeMiA and Industry Contributions to Testing Automation - JAMAICA 2013.

[11] Razak, R. A., & Fahrurazi, F. R. (2011). Agile testing with Selenium. 2011 Malaysian Conference in Software Engineering.

[12] Jain, R., Johnson, B. K., & Hess, H. L. (2015, July). Performance of line protection and supervisory elements for doubly fed wind turbines. In 2015 IEEE Power & Energy Society General Meeting (pp. 1-5). IEEE.

[13] Bures, M., & Filipsky, M. (2016). SmartDriver: Extension of selenium WebDriver to create more efficient automated tests. In 2016 6th International Conference on IT Convergence and Security (ICITCS) (pp. 1-4). IEEE

[14] Srinivas, K., & Prakash, C. (2017). A Comparative Study of Testing Framework with Special Emphasis on Selenium for Financial Applications. International Journal of Soft Computing, 12(3), 148-155.

[15] Hanna, M., Aboutabl, A. E., & Mostafa, M. S. M. (2018). Automated Software Testing Framework for Web Applications. International Journal of Applied Engineering Research, 13(11), 9758-9767.

[16] Sneha, K., & Malle, G. M. (2017). Research on software testing techniques and software automation testing tools. In 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS) (pp. 77-81). IEEE.

[17] Dhir, S., & Kumar, D. (2019). Automation Software Testing on Web-Based Application. In Software Engineering (pp. 691-698). Springer, Singapore.

[18] Berłowski, J., Chruściel, P., Kasprzyk, M., Konaniec, I., & Jureczko, M. (2016). Highly automated agile testing process: An industrial case study. e-Informatica Software Engineering Journal, 10(1).

[19] Ramya, P., Sindhura, V., & Sagar, P. V. (2017). Testing using selenium web driver. 2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT).

[20] Singh, I., & Tarika, B. (2014). Comparative analysis of open source automated software testing tools: Selenium, sikuli and watir. International Journal of Information & Computation Technology, 4(15), 1507-1518.

[21] Stocco, A., Leotta, M., Ricca, F., & Tonella, P. (2016). APOGEN: automatic page object generator for web testing. Software Quality Journal, 25(3), 1007–1039.

[22] Mahalakshmi, M., & Sundararajan, M. (2013). Traditional SDLC Vs Scrum Methodology–A Comparative Study. International Journal of Emerging Technology and Advanced Engineering, 3(6), 192-196.

[23] Petre, I., Smada, D. M., & Boncea, R. (2017). A user-centric approach to test automation of web-based applications. In RoCHI (pp. 49-52).

[24] VersionOne. (2018). 12th Annual State of Agile Report.

AUTHOR

Vijay is a qualified Software Test Engineer and an engineering graduate in Information and Technology from AGI institute in Graduate Diploma in Software Testing and have 5 years of experience in software testing competency. Vijay have been working in Cognizant Technology Solutions for around 3 years. After completion of the graduation course Vijay is currently working in Duco Consultancy as a Test Engineer in New Zealand.

Dr. Shahid Ali is a senior lecturer and IT program leader at AGI Education Limited, Auckland, New Zealand. He has published number of research papers on ensemble learning. His expertise and research interests include ensemble learning, machine learning, data mining and knowledge discovery.

.

CRITICAL DISCOURSE ANALYSIS OF HUAWEI ON TWITTER

Zheng qiqi

Guangdong University of Foreign Studies, China

ABSTRACT

Comparing with traditional media, online social media seem to provide more opportunities for people to speak out their ideas. Twitter integrates the whole world users in the platform, so that users from different places can exchange their views in the Internet. These functions own the decentralized feature, which is expected to change the original power structure in international communication. Huawei is an outstanding representative of Chinese company whose texts to some extent illustrate the overseas public 's evaluation of Chinese image.

On that basis, this study adopts the theory of Fairclough's Critical Discourse Analysis and analyzes the ways of Huawei discourse on Twitter. In this way, the current paper tries to investigate the production, distribution and consumption of Huawei discourse on Twitter. Meanwhile, this dissertation also attempts to discuss the situation of the construction and dissolution of the power structure behind social media in the new media era.

KEYWORDS

Twitter, Huawei, CDA theory, Mengwanzhou, Renzhengfei

1. INTRODUCTION

It is widely acknowledged in current society that people are easier to receive information from all different others around the word compared to the past. Especially nowadays, thanks to the Internet, some online media such as Facebook and twitter can display users' ideas from every corner of the word. Additionally, social media seems to give everyone right to express their thoughts equally because of decentralized network.

However, reviewing literature about mass media reveals that there is always existing power construe behind the media. Fairclough (2003) argues that ideology is a complicated situation in the media discourse for it is flexible in the discourse. Therefore, it is important to investigate whether ideology and power construe exist in social media in a theoretical analysis. On that basis, a theory is needed for the investigation of power and ideology behind mass media. As a result, this paper adapts the Critical Discourse Analysis (CDA) framework, which was came up by Fairclough. He(Fairclough ,2003) emphasizes the media discourse research could divided into three dimensions. The first dimension is text analysis; the second dimension considers discursive practice, while the third-dimension regards social practice as a deeper study. This article aims to study how the image of Chinese high technology company Huawei was represented through the discourse of international social media Twitter.

Various values are shaped by the mass media. Fairclough (2003) stressed that the institutions such as newspaper, spread the ideology by using the information to influence the social view. The purpose of the research is to explore the ideology and power behind Twitter by applying CDA theory to Huawei discourse on Twitter.

David C. Wyld et al. (Eds): SPTM, IPPR, AIS, CSIT, DaMi, AMLA - 2020
pp. 115-123, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100611

The context behind this paper is a series of events about Huawei. In December of 2018, Meng Wanzhou was arrested in Canada for connecting business dealing with Iran. Meng is chief financial officer of Huawei, China's largest telecommunications supplier（Kim,2018）. In 2019, the US government baned Huawei for national security, which was accused by China that Washington tried to halt the rise of Chinese high technology company (Sean,2020).

Although this heat debt still keeps going, the events arise plenty of attentions all the world. The ideological divides between China and America are obvious, so that ideology could be exhibited in discussion on Twitter. The objectives of the present study are to trace power relations and ideology through the users' text in the certain online platform. The following part will explain the methodology of theory and way of collecting the data. Following that, the data will be analyzed in three dimensions with third part, and after that the final part is the conclusion of finding of Huawei discourse on Twitter.

2. METHODOLOGY

2.1. Theoretical consideration

This paper accepts Fairclough's model of CDA involving three dimensions, text, discursive discourse and social practices to analysis media. However, when CDA theory was presented to analysis newspapers and television in 1996, some deeper investigations are not suitable in social media nowadays(Hart, 2013). As a result, this part will explain the specific theoretical consideration in three dimensions.

Text analysis, according to Fairclough, there are four main ways. Text on twitter are made by various users. So it is not fixed habit of grammar, cohesion, and text structure (Hoffmann et al., 2017). Therefore, the deeper investigations of text analysis on twitter is language.

Furthermore, discursive discourse under three more sections, the production, distribution and consumption of discourse. Therefore, the discourse of twitter could be examined in intertextuality. It is clear to reflect the sources and producer of discourse under examine the intertextuality (Zinken, 2003). Intertextuality refers to the text attribute of "the performance of other texts in any text", which is formally put forward by Kristeva (1980). After that, the concept of intertextuality was applied to many fields such as linguistics and communication by many scholars. Intertextuality has various definitions in history. In broad terms, it refers to the influence relationship of multiple discourses reflected in the text (Martin, 2011). In a narrow sense, it refers to the content of A text being presented in B text in some way (Zengin, 2016). Among them, the French scholar Samoyault(2003) explained the intertextuality in detail in his book "Intertextuality Research", including the classification of intertextual techniques and intertextual behaviors that link to different attitudes of discourse makers.

For this reason, in the production of discourse, the paper will classify texts that refer to the content of other users as "inter-text of intertextuality ", that is, inter-text within text. In terms of text distribution and consumption, this research classifies the disseminated texts of other users' texts as "extra-text of intertextuality ", that is, intertextualities other than utterance texts.

As for social practices, Fairclough（2003）considers power as a core concept of CDA, because power relations sharp the value and hegemony directly influencing the discourse. So the research takes power relation as standpoint to analyze the phenomenon.

2.2. A Selection of Data

Based on the CDA model, this paper collects the texts from Twitter to have a further analysis. The specific steps of this article are following. The author searched the keywords # huawei #, # renzhengfei # and # mengwanzhou # on Twitter to obtain 4,776 texts from December 1, 2018 to December 31, 2019. According to the equidistant sampling, 981 original samples were randomly selected. Later, in order to obtain more comprehensive user information, the author cleans the sample data, and performs manual statistics and translation on the user information address. This is because the Twitter user address is not automatically generated based on the user's location, but is freely written by the user. Some users do not set the address, or although set the content, but it is the non-existence address, such as leakland, which are fake addresses.

According to the data collection, the author also makes statistics on the user's social identity, and divides them into two categories: institutional and individual users. The organization is divided into three categories: news organizations, technology website organizations and Huawei official. In terms of individual users, although opinion leaders and average users are individual users, their followers and influence on Twitter are different. More influence means that their words are read by more users compared to average users, so that the opinion leaders get more power in the discourse. According to the actual situation, there are two types of opinion leaders: one is a user who is authenticated by Twitter, and the other is a user who is not authenticated but has more than 10,000 followers. Two Twitter authentication methods are adapted in this part. One is that the platform uses algorithms to actively confirm the true identity of celebrities, and the other is that the platform passively audits the authentication applied by the user, and the probability of passing the audit is linked to the user's popularity. Therefore, the author believes that users authenticated by Twitter have certain popularity and can be regarded as opinion leaders. On the other hand, some users with more than 10,000 followers, (Parmelee et al., 2013), although authenticated, have received a high degree of attention, so they are classified as opinion leaders. In summary, there are five categories of social identity: news organizations, technology website organizations, Huawei officials, opinion leaders, and average users.

3. RESULT AND DISCUSSION

3.1. Texts

Metaphor is a vital investigation in language that is under the first dimension of CDA. Metaphor is a concept proposed by scholar Roman Jakobson (1990), which has always been regarded as a literary language. But Fairclough (2003) believes that it exists in all kinds of discourses. He thinks that when a specific metaphor is used to represent ideology.It is a specific way to recognize and construct reality, so metaphor is an important indicator of discourse analysis.

Metaphor refers to substituting metaphor for ontology, based on the similarities or categories proposed between the real subject and its metaphorical pronouns, to construct an alternative "reality" (Andreotti et al., 2011). The theme of Huawei on twitter also includes "reality" constructed by metaphors. When different users of discourse make some of the same metaphors and put them into their discourse, they construct "reality" from a new perspective. Behind these metaphors are also hidden positions and value judgments. There is Table 1.

Table 1. Use of Twitter "Huawei Discourse" Metaphor

Ontology	Metaphor
Huawei	Dirty bomb
	A security nightmare
	Claws of the Red Dragon
	PLA-er (People's Liberation Army-er)
	A company Full of thieving parasites
	A branch of Chinese Intelligence telecommunications company
	Trojan Horse
	Chinese telecoms giant 5G role
	National champion’s telecom gear
China	Communist China
	Nazi China
	Totalitarian state
Google-huawei Cooperation	Down the Rabbit hole

From the above table, in the discourses of different users, Huawei is constructed as a telecom giant who steals users’ information and attacks national security. Huawei not only has contacted with the Chinese government, but also helps African countries' dictatorships to monitor their political opponents and builds surveillance technology. Therefore, a huge security risk could be seen. These metaphors reflect the vigilance and hostility of such discourse users towards Huawei and even China.

One of the reason why these discourses occur is that China is always considered as a totalitarian state in media (Lee, 2020). Meanwhile, Huawei, as a Chinese company, is linked to China. So, in some metaphor, China uses Huawei as a weapon to attack the national security of other countries, and even spread and penetrate communist viruses. This is a kind of mind of zero-sum game, and it is obviously the Cold War mentality, and it is the confrontation between different ideologies behind the texts.

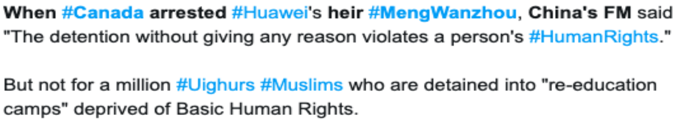
Through the use of metaphors, some discourse users construct Huawei in an incorrect or even distorted "reality", which reflects the disparate position and ideology of the discourse producers.

3.2. Discursive discourse

The production, distribution and consumption of discourse are three elements in discursive discourse analysis of CDA theory. This part analyses the production, distribution and consumption of discourse by tracing howthe Tweets created and who consumed them.

Tweets about topicsof Huawei could be presented in other people's tweets in various ways on Twitter. This way will form a model of a single text with multiple creators. These methods include direct citation, indirect citation, retweet and reference.

Direct quotation is a direct use of some words by the text producer, usually quoted by quotation marks.

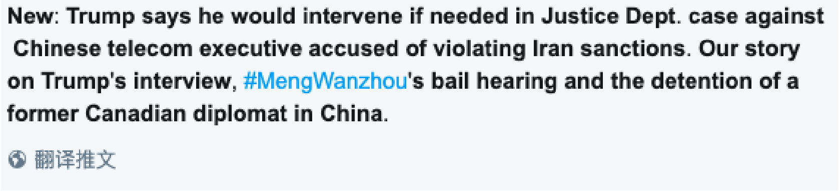


When #Canada arrested #Huawei's heir #MengWanzhou, China's FM said "The detention without giving any reason violates a person's #HumanRights."

But not for a million #Uighurs #Muslims who are detained into "re-education camps" deprived of Basic Human Rights.

Figure1. direct citation

Although indirect quotation is a quotation in dialogue, it does not use quotation marks, but marks the original speaker to achieve the quotation effect by indicating the speaker.

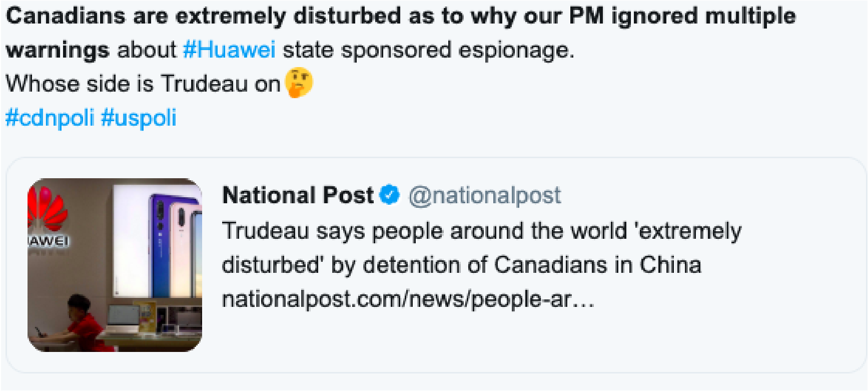


New: Trump says he would intervene if needed in Justice Dept. case against Chinese telecom executive accused of violating Iran sanctions. Our story on Trump's interview, #MengWanzhou's bail hearing and the detention of a former Canadian diplomat in China.

翻译推文

Figure2. Indirect citation.

Retweet refers to forwarding the content of other users' tweets, such as speeches or website articles, to their own public homepage. This is the most common speaking behavior on Twitter. The content quoted in the retweet will be marked below the Tweet. Apart from that, the content of the quoted article or tweet is visible, and users can see the full content by clicking it.



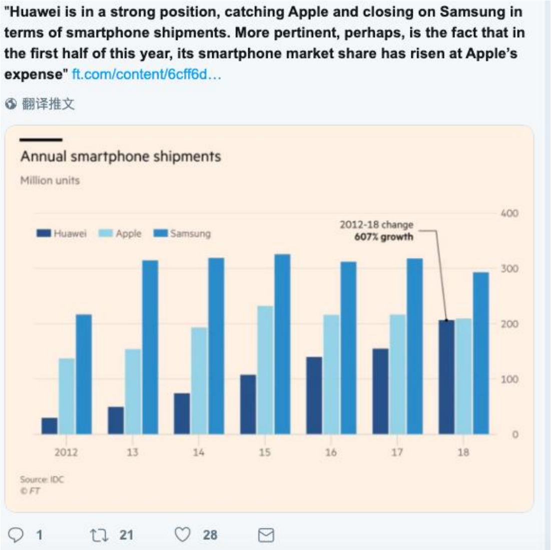
Canadians are extremely disturbed as to why our PM ignored multiple warnings about #Huawei state sponsored espionage.
Whose side is Trudeau on 🤔
#cdnpoli #uspoli

National Post @nationalpost
Trudeau says people around the world 'extremely disturbed' by detention of Canadians in China
nationalpost.com/news/people-ar…

Figure 3. Retweet

There are two types of reference classifications as Samoyault (2003) mentioned. One is accurate reference materials, such as literature citations of academic works, and the other is simple references. The following analysis adopts a simple reference standard, that is, the mentioned topics can be traced back to the text.

In Twitter, the links attached to the text are mostly simple references. For example, some producers put a reference article link at the end of the text, and the content of the link has a certain relationship with the text itself, as shown in the following figure.



"Huawei is in a strong position, catching Apple and closing on Samsung in terms of smartphone shipments. More pertinent, perhaps, is the fact that in the first half of this year, its smartphone market share has risen at Apple's expense" ft.com/content/6cff6d…

Figure 4. Reference

All of types could be considered as “pretext”, which means the contents are citied by the others. Therefore, by discovering the categories of how the users generate texts, it is clear to find out what kind of texts will affect the discourse.

From the statistics, it can be seen that it is the organization that has the dominant position in the power structure of this link, and the individual producer has become the dominant person.

Table 2. Social identification of pre-text

Social identification	news organizations	technology website organizations	Huawei officials	Average users	opinion leaders
percentage	75.8%	9.6%	4.8%	1%	8.8%

In the production of discourse, more than 50% of the previous texts quoted by its content are from institutions. This shows that the text from the institution maybe a main entrance of information for the discourse makers. In the other words, institutions control the production of Huawei discourse on Twitter.

When it comes to the distribution and consumption of discourse, these parts need to be choose indicators that can reflect the communication paths and consumption texts because it has to reflect the audience's communication and interpretation of the text. However, since the most direct data "tweet page views" cannot be visually observed, "retweet number" was chosen in this study, which could directly reflect the communication path and the consumption text.

Retweet function on Twitter will show the dialogue in users’ homepage, because both parties can be presented on the platform when they comment. Therefore, on the Twitter platform, discourse production is text creation based on other texts, and discourse consumption refers to interactive actions such as commenting on others’ texts. As a result, retweet is a kind of extra-text of intertextuality that directly refers to the consumption of discourse.

Table 3. Social identification of retweet users

Social identification	Average users	Opinion leader	News organization	Non-identification
percentage	86.4%	5.7%	1.3%	6.6%

In the distribution and consumption of Huawei discourse, more than 80% of the retweets are average individual users. Therefore, it could be stated that it is the individual users who dominate the distribution and consumption of Huawei discourse.

To be specific, the average individual users are the dominant part in the consumption of discourse. They occupy the leading role in continued production in consumption.

However, no matter what kinds of quotation, the original text has been reproduced by the users, so it probably could be influenced by users' ideology. The first type of direct quotation using quotation marks can make the original discourse appear obvious, but the discourse text of the "markup" is still selected by the discourse creator (Maier, 2015). Therefore, the discourse writer will use the selected discourse to justify his point of view, thereby indirectly imposing it on the audience. Indirect discourse is more conducive for creators to rewrite the words and intonation of sentences, which means the indirect citation also reflects the attitude of the producer of the discourse (Dijk, 1993). The form of retweet can be considered as fully independent. References are less involved in the construction of discourse. So, it is also important to know whether the information of pre-text influences the users.

In twitter, users could show their opinions of pre-text directly in their own texts. Based on that, the researcher considers there are generally two kinds of attitude to the pre-text, agree or disagree. From these attitude, the deeper investigation could show whether the opinions of institution will be accepted by the users who quote.

Table 4. The attitude to pre-text

	agree	disagree
Percentage	79.2%	20.8%

3.3. Social Practice

The power contracture may change due to different social identities, relationships and other factors (Fairclough,2003). But when a power plays a leading role in a class of discourse, it could be determined that this kind of power occupies a certain discourse power or discourse advantage (Guo,2019). It at least shows two totally different power relations in the discourse.

Although the data shows that the participants of Huawei's discourse are all over the world and their identities are diverse, there is still a power structure behind the production, distribution and consumption of discourse.

From the statistical data above, first of all, news organizations and institution control the discourse production but behind the text. The institution does not directly control the discourse by news report but the information recourses. Once users on Twitter quote their information, it is easy for them to trap in the value already set up by the institution.

Secondly, the power relation of discourse consumption emphasizes the superior of average users. In exploring the discourse distribution and consumption parts, no matter what kinds of discourse inclination, more than 80% of the retweet users' identity data are average users, and other types of social identity users account for this proportion very little. Therefore, it can be explained that average users have become the absolute main force in this power relation, forming a power structure dominated by average users. The reason is the decentralized network technology provide average users with the possibility of contributing a new power structure.

Twitter as a social media, bring all the users in the same platform and give them the same opportunity to speak their own ideas. The decentralized features allow people to support the texts they like. So, they control the consumption of the discourse.

4. CONCLUSION

In the current international communication research, whether it is Dependency theory, Cultural Imperialism theory, Media Imperialism theory or World system theory, it is believed that the current international communication pattern presents the phenomenon that the national media of developed countries dominates（Ming,2006）. However, it seems that the social media that owns the decentralized technology bring some changes to this power relation.

Fairclough's CDA is useful in investigating the social discourse. Applying the theory to Huawei discourse could have deeper learning in the power structure behind the discourse. So, the article examined the three dimensions of Huawei discourse on Twitter from the perspective of CDA theory, and found that institutions, especially news institutions, still control the production of the previous text, and decentralized social media has not shaken this power structure.

However, average users actively participate in the discourse consumption link, thereby forming a dominant force in this link and dissolving some of the media power. It can be seen that although the current online social media has not shaken the original media power, the social media with decentralized technology partly dissolve the power structure.

This study still has some drawbacks, such as the accuracy of social identification. In the future, we can try to have more accurate information and analyze the discourse more variables. In addition, this study could not consider the fake followers.For example, fake follower accounts are controlled by robotsand they are used to get averifiedsymbol (Cresciab et al., 2015).

Some users would buy Twitter followers to pretend the opinion leader (Stringhini,2013). This inaccurate information will influence the data. Besides,Twitter restrict the data mining in tweetpage views, which means no direct data of the consumption and margin of error probably be included in result (Aral &amp; Zhao, 2019). So, thefuture study may be useful for sample surveying proper margin of error, but the large sample may need to consider another method.

ACKNOWLEDGEMENTS

I would like to thank everyone.

REFERENCES

[1] Fairclough Norman.&Yin, X. Hua Yu She Hui Bian Qian, Beijing: Hua xia chu ban she.2003

[2] H, Kim. Huawei CFO Meng Wanzhou Arrest: China: US Trying To 'Stifle' Its Rise. [online] USAtoday. Available at: <https://www.usatoday.com/story/news/world/2018/12/07/huawei-cfo-meng-wanzhou-arrest-china-says-u-s-trying-stifle-its-rise/2236365002/> [Accessed 14 May 2020].

[3] K, Sean. Huawei Ban Timeline As Trump Extends Executive Order Targeting Chinese Company. [online] CNET. Available at: <https://www.cnet.com/news/huawei-ban-full-timeline-us-restrictions-china-trump-commerce-department-security-threat-5g-p40-coronavirus/> [Accessed 14 May 2020].

[4] C. Hart, “Event-construal in press reports of violence in two recent political protests,” Journal of Language and Politics, vol. 12, no. 3, pp. 400–423, 2013.

[5] C. R. Hoffmann and W. Bublitz, “12. Discourse and cohesion,” in Pragmatics of Social Media, Berlin: De Gruyter Mouton, 2017, pp. 317–349.

[6] J. Zinken, “Ideological Imagination: Intertextual and Correlational Metaphors in Political Discourse,” Discourse & Society, vol. 14, no. 4, pp. 507–523, 2003.

[7] J, Kristeva. Desire In Language:A Semiotic Approach To Literature And Art, New York: Columbia University Press. 1980

[8] E. Martin, “Intertextuality: An Introduction,” The Comparatist, vol. 35, no. 1, pp. 148–151, 2011.

[9] M. Zengin, “An Introduction to Intertextuality as a Literary Theory: Definitions, Axioms and the Originators,” Pamukkale University Journal of Social Sciences Institute, vol. 2016, no. 50, pp. 299–327, 2016.

[10] D Samoyault, S, Wei. Hu Wen Xing Yan Jiu, Network Books, Tianjin: Tianjin ren min chu ban she.2003

[11] J. H. Parmelee and S. L. Bichard, “Introduction：the important of Twitter of polities,” in Politics and the Twitter revolution: how tweets influence the relationship between political leaders and the public, Lanham, Mar.: Lexington Books, 2013, pp. 1–35.

[12] Jakobson, R., Pomorska, K.&Rudy, S.Language in Literature, Cambridge: The Belknap Press of Harvard University Press. 1990

[13] V. Andreotti, C. Ahenakew, and G. Cooper, “Epistemological Pluralism,” AlterNative: An International Journal of Indigenous Peoples, vol. 7, no. 1, pp. 40–50, 2011.

[14] C.-C. Lee, CHINA'S MEDIA, MEDIA'S CHINA. S.l.: ROUTLEDGE, 2020.

[15] E. Maier, “Quotation and Unquotation in Free Indirect Discourse,” Mind & Language, vol. 30, no. 3, pp. 345–373, 2015.

[16] T. A. V. Dijk, “Principles of Critical Discourse Analysis,” Discourse & Society, vol. 4, no. 2, pp. 249–283, 1993.

[17] J, Guo. The Discourse Construction of the Overseas Social Platform "Belt and Road", Guangzhou: Guangdong University of Foreign Studies. 2019

[18] M, An.“On the Global Communication Pattern in the New Century“, Modern Communication (Journal of Communication University of China), vol 06, pp 20-24, 2006

[19] G, Stringhini and G, Wang, “Follow the green: growth and dynamics in twitter follower markets,” in Porc, IMC '13. 2013. pp.163–176.

[20] S. C. Cresciab, R. D. Pietro, and M. Petrocchi, “Fame for sale: Efficient detection of fake Twitter followers,” Decision Support Systems, vol. 80, pp. 56–71, Dec. 2015.

[21] S. Aral and M. Zhao, “Social Media Sharing and Online News Consumption,” SSRN Electronic Journal, 2019.

AUTHOR

Zheng, qiqi Short Biography.

AUTHOR INDEX