ISSN: 2581-3242

Universidad de Oviedo

CMR INSTITUTE OF TECHNOLOGY EXPLORE TO INVENT

International Journal of MACHINE LEARNING AND NETWORKED COLLABORATIVE ENGINEERING

Vol.2 No.2 2018

PREFACE

DOI: https://doi.org/10.30991/IJMLNCE.2018v02i02

The International Journal of Machine Learning and Networked Collaborative Engineering (IJMLNCE) is a quarterly published, open access, peer-reviewed, international journal with ISSN No 2581-3242**.** After one year of journey, we achieved a milestone, after publishing the first three editions. The Journal is now indexed with some popular indexing like BASE, CrossRef, CiteFactor, DRJI, Google Scholar, Index Copernicus, J-Gate, PKP-Index, ROAD, Scilit, and Socolar.

Now, the fourth edition is very close to be online, in the form of Volume No-02 Issue No-02. For this particular issue of volume, we have five innovative and outstanding papers, within the scope of the journal, covering different aspects of machine learning and collaborative engineering, which are highly emerging research area nowadays.

The paper, "Steganalysis for Reversible Data Hiding based on Neural Networks and Convolutional Neural Networks", by Huong Thom et al. focuses on the innovative approach of neural networks in case of steganalysis for reversible data hiding. The aim of the paper is to restore the original images after extracting information from a hidden image with secret data. Their research work emphasizes on the method to improve the detection rate of images with 96% correct detection rates using neural networks and 94% with convolutional neural networks [1].

In the next paper, "Prediction Model for Pollutants with Onboard Diagnostic Sensors in Vehicles" by Maldonado et al., focuses on a prediction model for pollutants with onboard diagnostic sensors in vehicles. The aim of this paper is to show the relationship between the internal parameters of on-road vehicles and their emissions. Internal values are collected through the On-Board Diagnostics port, while values of the emissions are measured from the exhaust pipe using an Arduino board. There are observable correlations between carbon dioxide emissions and vehicle speed, as well as carbon dioxide emissions and engine revolutions per minute [2].

The next paper "Information Processing in Neuron with Exponential Distributed Delay", by Choudhary et al. is also from the field of neural network. This paper aims to show how to use neural networks for information processing. This paper has been claimed as an extension work of the Leaky Integrate-and-Fire model, analyzing the impact of the exponentially distributed delay memory kernel on spiking activity and steady-state membrane potential distribution. In this paper, the model has been implemented using recurrent neural networks, which show more accuracy and a potential way to implement chip level artificial intelligence [3].

The next paper, "Improvement of Automated Learning Methods based on Linear Learning Algorithms", by Kiani, F., chosen from the automated learning area. The author proposes a new approach to improve automated learning methods based on the reinforcement learning technique. The effectiveness of the interactions with the environment is evaluated by the number of rewards and penalties that are taken from it. Kiani proposes three versions: simple, sequential and unstructured linear learning methods, that focus on different scenarios and areas [4].

The last paper of this issue is "IoT with Big Data Framework using Machine Learning Approach", by Chatterjee, J. M. focuses on a Big Data framework collaborated with the IoT, using machine learning. In this paper, the author gives an explanation about the relationship between big data and the Internet of Things, together with different issues and challenges, all from the point of view of offering solutions using an approach based on machine learning strategies [5].

REFERENCES

[1]. Huong Thom, H. T., Anh, N. K., & Vu, B. D. (2018). Steganalysis for Reversible Data Hiding Based on Neural Networks and Convolutional Neural Networks. International Journal of Machine Learning and Networked Collaborative Engineering, 2(2), 40-48.

https://doi.org/10.30991/IJMLNCE.2018v02i02.001

[2]. Maldonado, B., Bennabi, M., García-Díaz, V., González García, C., &Núñez-Valdez, E. R. (2018). Prediction Model for Pollutants with Onboard Diagnostic Sensors in

Vehicles. International Journal of Machine Learning and Networked Collaborative Engineering, 2(2), 49-57.

https://doi.org/10.30991/IJMLNCE.2018v02i02.002

[3]. Choudhary, S. K., & Bharti, S. K. (2018). Information Processing in Neuron with Exponential Distributed Delay. International Journal of Machine Learning and Networked Collaborative Engineering, 2(2), 58-66.

https://doi.org/10.30991/IJMLNCE.2018v02i02.003

[4]. Kiani, F. (2018). Improvement of Automated Learning Methods based on Linear Learning Algorithms. International Journal of Machine Learning and Networked Collaborative Engineering, 2(2), 67-74.

https://doi.org/10.30991/IJMLNCE.2018v02i02.004

[5].Chatterjee, J. M. (2018). IoT with Big Data Framework using Machine Learning Approach. International Journal of Machine Learning and Networked Collaborative Engineering, 2(2), 75-85. https://doi.org/10.30991/IJMLNCE.2018v02i02.005

Editor-in-Chief

International Journal of Machine Learning and Networked Collaborative

Vicente García-Díaz and Vijender Kumar Solanki

EDITORIAL BOARD MEMBER

EDITORIAL BOARD MEMBER

INDEX

EDITORIAL OFFICE

CALL FOR PAPER

International Journal of Machine Learning and Networked Collaborative Engineering (IJMLNCE) with ISSN **2581-3242**, is a quarterly published, open access, peer-reviewed, international journal, focuses on publishing authentic and unpublished quality research papers in the emerging field of Machine Learning and Collaborative Engineering.

This is a scientific journal aims to promote highly qualitative and innovative research works, not limited to but focuses on the area of machine learning, collaborative engineering, and allied areas. We publish only original research articles, review articles, and technical notes.

The journal is registered with CrossRef **10.30991/IJMLNCE**. Each accepted manuscript shall be assigned a unique digital object identifier for identification. The abbreviated key title for our journal allocated by ISSN is **Int. j. mach. learn. Networkedcollab. eng**.

The journal appears in popular indexes like **BASE (Bielefeld Academic Search Engine), CrossRef, CiteFactor, DRJI, Google Scholar, Index Copernicus, J-Gate, Portico, PKP-Index, ROAD, Scilit, Socolar**. Our journals manuscripts can also be found in the directory of **ReserachGate & Kudos**.

In this present technological era, the areas like machine intelligence, machine learning, and its associated domains are one of the most popular and demanding choices for the researchers as well as the industry personnel.

In last few years, numerous uses of machine learning and its related domain, has drawn ample attention of the people, that has generated a large number of applications in this field, making machine learning and collaborative engineering highly admired one.

Machine intelligence or machine learning is not a new concept. In terms of Artificial Intelligence, we were familiar with several aspects of the field, but nowadays with the introduction of machine learning, the use of this has been highly evolving, especially for improving the lifestyle of the human being.

There are numerous application areas of machine learning or machine intelligence, irrespective of any famous sector, from Healthcare, Space, Automation, Aviation industries etc. to entertainment industry and even in academia.

With this insight, IJMLNCE invites quality and authentic research works in form of manuscripts, for our forthcoming version. The Journal solicits latest research work carried out by the researchers, academicians and industry practitioners, in the domain of machine learning and collaborative engineering.

Plagiarism is considered as a serious ethical and professional malpractice in case of IJMLNCE. So similarity check should be taken care at its highest level. The originality of the work will be highly appreciated. Each paper will be peer-reviewed by the expert reviewers and in a multilayer approach. The editorial board of the journal believes in high moral and ethical practices.

The authors are invited to submit their authentic and unpublished manuscripts online only. They will get final decision about the acceptance or rejection of the manuscript within 06-12 weeks from the date of submission of manuscripts. Once the manuscript gets accepted, author will be asked to complete the copyright agreement. An e-mail communication will also be sent to the author's registered Mail ID, after verification of incorporation of all reviewers' comments, in term of their final camera-ready-paper; it will be published to the upcoming journal's volumes.

Please note that there is **no publication fee**; **no article processing cost** for publishing articles in this journal and the journal is open access.

Vol. 02, No. 2 ,(2018), 040–048, **ISSN:2581-3242**

IJMLNCE JOURNAL

International Journal of Machine Learning and Networked Collaborative Engineering

Journal Homepage: http://www.mlnce.net/home/index.html

DOI : https://doi.org/10.30991/IJMLNCE.2018v02i02.001

Steganalysis for Reversible Data Hiding Based on Neural Networks and Convolutional Neural Networks

[a]Ho Thi Huong Thom*, [b]Nguyen Kim Anh, [c]Bui Dinh Vu

[a,b,c]*Faculty of Information Technology, Vietnam Maritime University, Vietnam*

[a]*thomhth@vimaru.edu.vn, https://orcid.org/0000-0002-1825-9577,*

[b]*Kimanh@vimaru.edu.vn, https://orcid.org/0000-0001-7725-7945*

[c]*buidv@vimaru.edu.vn,https://orcid.org/0000-0002-2914-9613*

Abstract

Lossless data hiding techniques is a technique that is very interesting. In which there is a large amount of reversible information hidden technologies. This technique makes it possible to restore the original image after extracting the information from the stego image. The stego image (hidden image with secret data) is hardly detected by any variable. There are many studies for this field are published. Secret information is hidden on the pixel space, frequency (cosine, wavelet) coefficient space or difference image coefficient space. However, by analyzing meticulously between the cover image and the stego image on these space, one can detect abnormal signs. In a previous work, a steganalytic techniques produced that was based on analysis of the transform coefficient histogram with the correct detection ratio between 88% and 92%. In this article, proposing another method to improve the detection ratio of that steganalysis based on Neural Networks (NNs) and Convolutional Neural Networks (CNNs). The test results show 96% correct detection rates for NNs and 94% for CNNs, this is a better result than our previous method. This proposed approach can be applied to detect stego images on spatial and other frequency domain.

Keywords

Steganography,

Steganalysis,

Cover Image,

Stego Image,

Histogram Shifting,

Lossless Data Hiding,

Neural Networks and

Convolutional Neural Networks

1 Introduction

Steganography is the science of hiding the existence of information. Its purpose is to convey the message secretly so except the sender and receiver, no one knows about the existence of the message. Steganography actually is a form of security by concealment. digital image, video, sound or any other files can be used as “cover” to carry secret messages, after that it is called stego.

* Corresponding Author

Ho Thi Huong Thom

Email: thomhth@vimaru.edu.vn

Steganalysis is opposite of Steganography. It is the science of detecting the message hidden using steganography. The main purpose task of it is to distinguish between cover image and stego image.

In recent years, many lossless data hiding techniques have been proposed for stego images. Lossless data can be embedded in the spatial domain [1, 2, 3] or in the transform domain [4, 5]. Xuan et al [5] proposed a method by a histogram shifting in integer wavelet transform domain (IWH method). This method hides messages into high frequent sub-bands of integer wavelet coefficients.

In [7], we offered a new steganalytic method based on integer wavelet transform that can detect stego images using Xuan's method. Besides capability of detecting the hidden image, the algorithm can estimate the length of embedded data reliably. To increase detection ratio, we also research another method based on Neural Networks (NNs) and Convolutional Neural Networks (CNNs). Recent CNN works focused on problems about computer vision, such as identification of 3D objects, natural images, and traffic signs [8, 9, 10] image denoising [11] and image segmentation [12]. Convolutional architectures also seem to benefit of unsupervised learning algorithms used for analyzing image data [13, 14, 15].

In the next section, we describe again the Xuan's steganography method. In section 3, 4 we review again our proposed steganalytic methods and introduce new method on NNs and CNN's. Our experimental results are displayed in section 5. Finally, the conclusion is given in section 6.

2 Lossless Data Hiding based on Integer Wavelet Histogram Shifting

In this section, we describe Xuan's IWH algorithm introduced in [5]. This algorithm does not cause distortion by hiding information on the integer wavelet domain. The image space domain after being transformed to integer wavelet domain will be divided into four sub-bands. Xuan et al hid secret information on three high-frequency domains. The details of the IWH algorithm are summarized as follows:

Suppose there are M bits of secret information into a high frequency band. The IWH algorithm performs the following steps:

Step 1: Choose the threshold T (T> 0), so that number of the coefficient in the range [-T, T] is greater than M. Set Peak = T.

Step 2: In the histogram of the wavelet coefficient, shift the histogram column (the value of the histogram is greater than the Peak) to the right by a unit to create a Zero column at the Peak +1 position. Secret information is hidden in this location. Scanning all the coefficients of the high frequency band, if the coefficient is equal to Peak, the secret bit is 1, add 1 to the Peak coefficient to become Peak + 1, the secret bit is 0, the value of coefficient doesn't change.

Step 3: Keeping the secret information, change Peak = - Peak, shift the histogram column (the value of histogram is smaller than Peak) to the left by one unit to create the Zero column at (-Peak – 1). Information is hidden at this point.

Step 4: If all M bits have been hidden, the algorithm stops here and records the stop position S = Peak. In contrast, set Peak = - Peak -1, go back to step 2 to continue to hide secret bits remains.

3 PreviousSteganalytic Methods

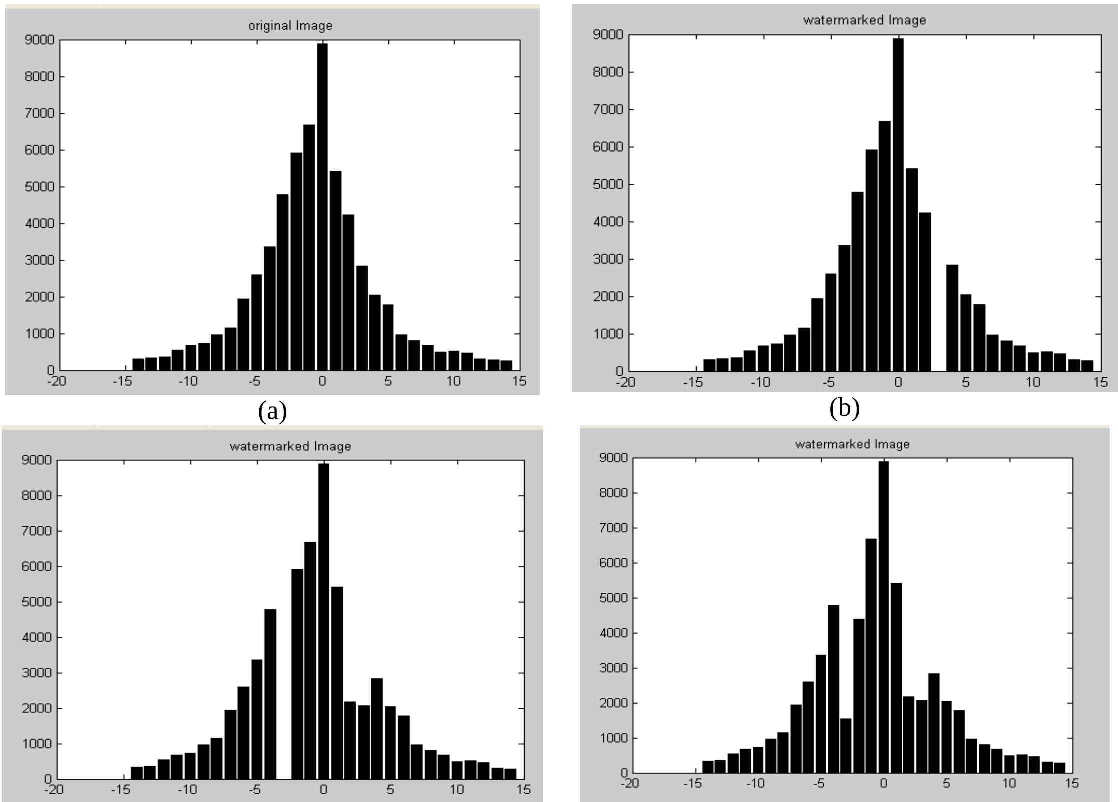
The IWH algorithm is a Lossless Data Hidingalgorithm, but if you compare the integer wavelet coefficients, the histogram of the original image and the IWH stego image, you will see this difference. This is the key to estimate the information hidden in the stego image.

We first give analysis of occurrences in watermarking process as the three following experiments:

In the first experiment, we use Lena image of size 512 x 512 pixels (see Fig. 1. (b)) and Logo image of size 128 x 56 pixels to test (see Fig. 1. (a)). After integer wavelet transform, we calculate the histograms of high frequency sub bands (see Fig. 2. (a)). We next embed payload data (that is the binary sequence from Logo image) into the high frequency sub bands with T=2 using IWH method. We get S= -2 and calculate again histogram of the high frequency sub bands that is shown in Fig. 2. (d). The data embedding process performs via some steps: the first and second step embeds data in the point 2 and -2 (see Fig.2. (b), (c)) but there are to been bedded data remaining, the process performs the third and fourth step with T= -2 to embed data (see Fig.2. (c), (d)).

a) b)

Fig.1: Test images: a) Lena original image, b) Binary logo image.



(a) (b)

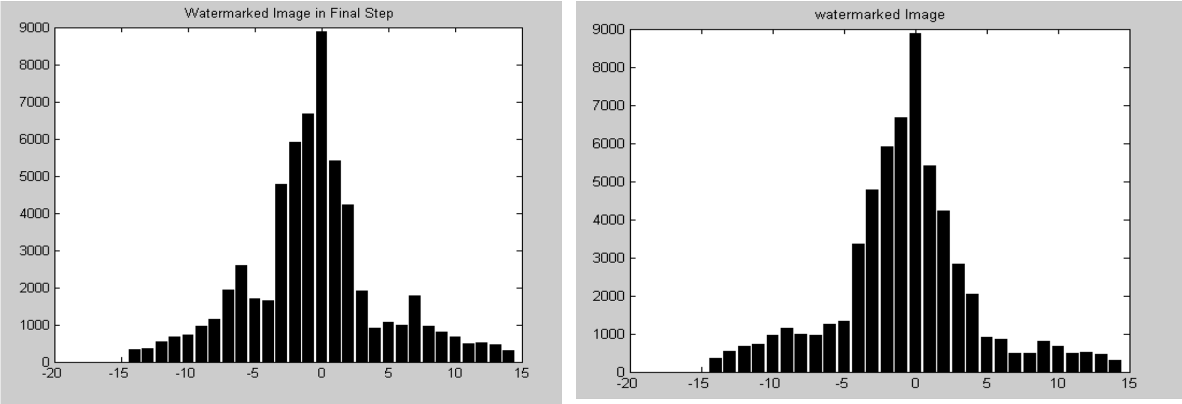
(c) (d)

Fig.2: An example showing how a zero point is generated and payload data embedding process: (a) original histogram, (b) histogram after a zero point is created, (c) histogram after data embedding at Peak =2 and then a new zero point is created at new next Peak, (d) histogram after data remaining embedding with new Peak.

In the second experiment, we use also the Lena original image and Logo watermark with T=4, we then get S=3. In this case, the histogram is changed much that is shown in Fig.3. (a).

In the third experiment, we use the same input with T=6, we then get S=-5. In this case, the histogram is changed clearly that is shown in Fig.3. (b).

From the above three experiments we find that the wavelet coefficient histogram is slightly variable. In the natural image, the value column of the pair of coefficients (h_i, h_{-i}) is approximately equal and symmetric across the value h_0 column, the h_i value is usually greater than h_{i-1} (Fig. 2. (a)). However, after hiding it broke the nature (Fig. 2. (d), Fig. 3. (a), (b)).



(a) (b)

Fig.3: Another example showing how payload data embedding process: (a) the histogram after data embedding with chosen T=4, (b) histogram after data embedding with chosen T=6.

From the analysis we can estimate length of the information be hidden in the stego images. Details of the algorithm show the following as:

Step 1: Set L=0 (length of message), scan all column h_i (i >0 and i <=max (all integer wavelet coefficient of high sub bands)), if the first $\frac{(h_i+h_{i+1})}{2} < h_{i+2}$ is met, stop scanning, set Peak = i be first location to estimate data length.

Step 2: if $h_{Peak} \approx h_{Peak+1}$, $L=L+h_{Peak}+h_{peak+1}$; set Peak = -Peak and perform next step 3. Contrariwise, perform step 4.

Step 3: if $h_{Peak} \approx h_{Peak+1}$, $L=L+h_{Peak}+h_{peak+1}$; set Peak = -Peak - 1 and return step 2. Otherwise, perform step 4.

Step 4: if $h_{Peak+1} < h_{Peak+2}$ and $h_{Peak+1} < h_{Peak}$ then $L= L+2*h_{Peak+1}$. The process stops here.

Applying the proposed algorithm for the three experiments we can estimate the length of information hidden in the image according to Table 1.

Table 1: Estimating the length of information hidden on Lena image

Embedded data length	Chosen threshold T	Gotten Stop value S	Estimated data length
7168	2	-2	7231
7168	4	3	6998
7168	6	-5	7177

4 Steganlysis based on Neural Networks

Base on the analysis in section 2, 3, we found out classify images by Artificial Neural Networks – NNs and Convolutional Neural Networks - CNNs.

On NNs, the same architecture is used for experiments on 1000 cover images which were downloaded from [16, 17] and 500 stego images from 1000 cover images. We transform all pixel of cover image and stego image to integer wavelet frequency domain, then calculate histogram of the wavelet coefficients $h=[h_{-128}, h_{-127}, \ldots, 0, \ldots, h_{126}, h_{127}]$, each h_i divided by max(h) to decrease h_i in value range [0,1] to increase the accuracy of the network training process. We consider the vector h is the characteristic vector as the input neuron (each h_i being an input neuron) for training the network. Using a hidden layer with 20 neurons, the neurons of the output layer are two neurons, using the sigmoid function to summary the output values of each class, the weights of the neurons in each class are initialized accordingly so that the smallest possible output error (Fig. 4).

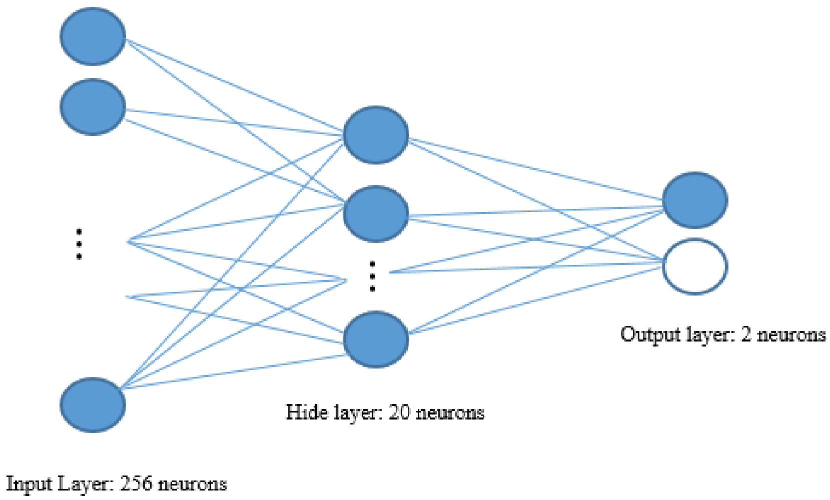


Fig.4: ANN architecture for the proposed method

On CNNs, use CNNs as base classifiers [11]. This network consists of 6 layers, 1 input layer, 2 convolution layers, 2 max – pooling layers and one layer of output. The input layer consists of 16 x 16 neurons (corresponding to 16 x 16 = 256 value of the wavelet coefficient histogram). The first hiding layer (the first convolution layer) consists of 6 maps of 12 x 12 neurons with 5 x 5 filter windows. The next hidden layer (max-pooling) is the output of the first convolution layer. Six mappings of 6 x 6 neurons use a 2 x 2 filter core.

The next convolution consists of 12 mappings of 6 x 6 neurons with 5 x 5 kernel. The second max - pooling layer consists of 12 mappings of 2 x 2 neurons using the 2 x 2 filter kernel. The output layer has a neuron per layer (corresponding to the original image layer and the stego image) (Fig. 5).We pick the trained CNN with the lowest validation error, and evaluate it on the corresponding test set.

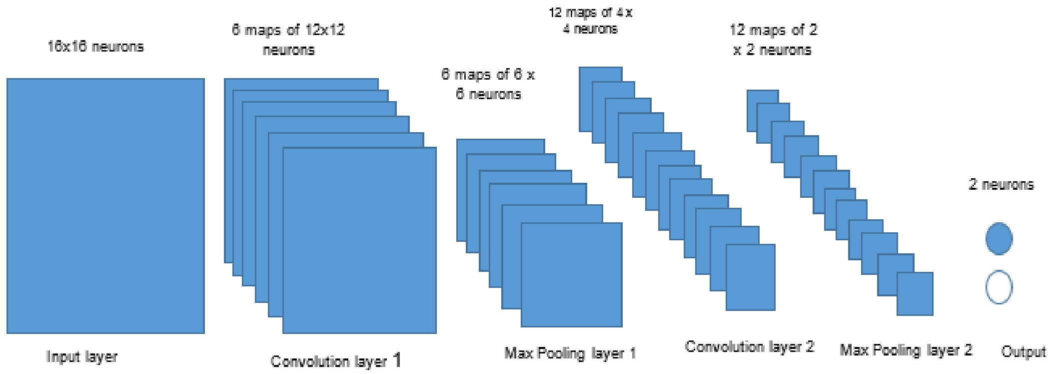


Fig. 5: CNN Structure for the proposed method

Prior to training, we give the detail of choosing input neurons of CNNs for the above steganographies: In a cover image or stego image in train dataset transforms integer wavelet transform, calculate the histograms h_w of high frequency sub bandswith values in set {-128… 127} (including 256 values of the wavelet coefficient histogram), divide h_D by max value of them to reduce the value 0 – 1, then reshape h into a vector of 16 x 16 elements. We use the vector as 16x16 input neurons of CNNs.

5 Experimental Results

Having a set of images, it includes 2088 images.They were downloaded from [15], [16] and they were created from my digital camera, all images are then converted to grayscale images by Photoshop CS2 software. This set of photos is created into two subsets they will be used to test the proposed detection or classification methods in above session as follows:

The first data set is used to detect by the estimation method includes:

+ 2088 cover images

+ 2088 stego images with hidden information is a 2000 bits binary string generated randomly.

The second data set is used for two neural networks includes:

+ The training set: 1500 images including 500 cover images and 1000 stego images which are embedded the randomly secret binary sequence of 2000 bits and 6000 bits into corresponding 500 original images by IWH method.

+ The testing set: 3676 cover images and stego images which are embedded the randomly secret binary sequence of 2000 bits or 6000 bits.

Proceeing to test three scenarios for two proposed approaches to detect hidden images using IWH hiding techniques:

+ Case 1: using estimating the information hidden in the wavelet coefficient domain of the image for the first data set, the images are hidden with the amount of information is 2000 bits, we obtain the estimation results as shown in Fig. 5. There the horizontal axis represents image number # and the vertical axis represents the embedded data length corresponding image number #.

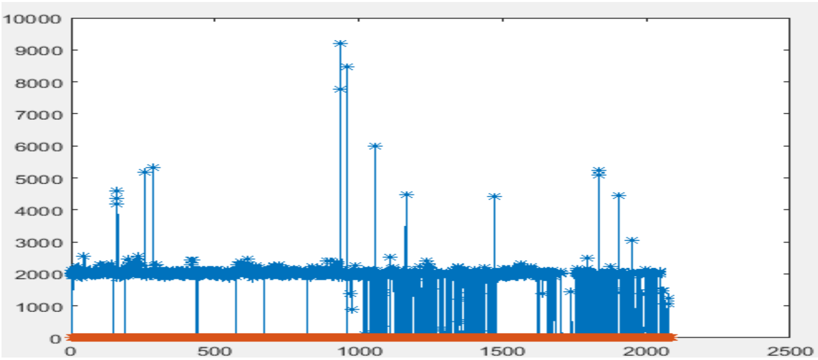


Fig.6: Estimate the hidden information for 2088 images with the hidden message of 2000 bits

Using estimating the information hidden in the wavelet coefficient domain of the image, the images are hidden with the amount of information is 6000 bits, we obtain the estimation results as shown in Fig. 6.

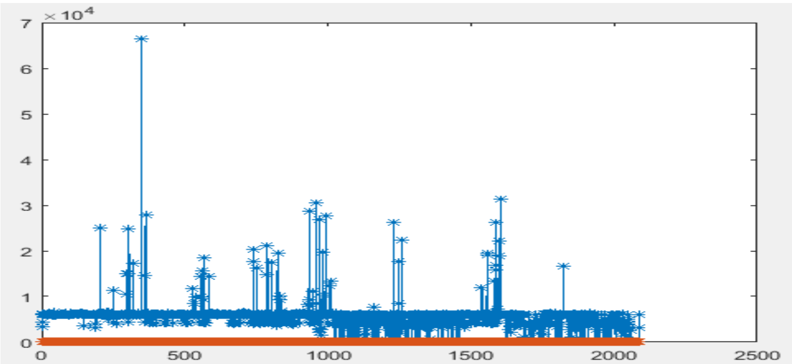


Fig. 7: Estimate the hidden information for 2088 images with the hidden message of 6000 bits

According to the first method the results are as follows:
- Original image (estimated information hidden by 0) - false detection of 440 image sat 21.07%
- Stego images with embedded 2000 bits (correct detection) rate reached 88. 93%, the average estimate is 1581.42 bits with a deviation of 933.39.The execution time is 147.43 seconds.

- Stego images with embedded 6000 bits (correct detection) rate reached 91.23%, the average estimate is 5301 bits with a deviation of 3165.The execution time is 156.65 seconds.

+ Case 2: using NNs to classify images for the second data set with epoch size options into training with 100 neurons (Number of batch), the results are as shown in the following table 2.

Table 2: The image classification resulting using NNs

Epoch	Training time (second)	classifying time (second)	True rate
10	0.27	0.19	94.31
50	1.332	0.22	96.32
100	2.81	0.18	96.21
200	5.37	0.2	95.72
300	8.12	0.2	95.51
500	14.14	0.2	95.23
Average		0.2	95.55

According to Table 2, the average detection rate for the data set is 95.55%, the classifying time is 0.2 seconds.

+ Case 3: using CNNs to classify images with epoch size options into training with 100 neurons (Number of batch), the results are as shown in the following table 3.

According to Table 3, the average detection rate for the data set is 93.05%, the classifying time is 0.38 seconds. However, the training time of CNNs is higher than the training time of NNs. From the three test cases, the results obtained by the method of message estimation are lower than those found in the Neuron network. In fact, CNNs are rated better than NNs, but in this case, NNs are better at detecting CNNs. This may be because the number of entry layers of CNNs in this case is much lower than those used in [11].In addition, the training time of CNNs is many times higher than that of NNs.

Table 3: The image classification resulting using CNNs

Epoch	Training time (second)	classifying time (second)	True rate
10	7.53	0.38	91.31
50	35.54	0.35	94.32
100	78.73	0.40	94.21
200	142.72	0.36	93.72
300	189.31	0.39	92.51
500	329.69	0.37	92.23
Average		0.38	93.05

6 Conclusions

In this paper, introducing two methods by using NNs and CNNs for better results. The two results show 96% correct detection rates for NNs and 94% for CNNs that indicates the reliability of the methods, this is a better result than old method [7]. Combining the old method [7] and new method to classify cover and stego images using IWH method. The first, using the new method (using on NNs or CNNs) to detect, then using the old method to estimate the hidden information.

However, it is hard to detect stego image with two factors which are shown in section 5. Noting that, there are many elements in this algorithms that can be changed or replaced with other elements. This research can be used to detect hidden images on spatial, frequency, or other domain.

References

[1]. O. Egger, M. Kunt, Embedded zerotree based lossless image coding, Proceedings., International Conferenceon Image Processing, Washington,DC,USA,USA,2015, https://doi.org/10.1109/icip.1995.537710

[2]. Ni, Z., Shi, Y., Ansari, N., Su, W.: Reversible data hiding. Proc. ISCAS (2003) 912–915. https://doi.org/10.1109/iscas.2003.1206123

[3]. Sang-Kwang Lee, Young-Ho Suh, and Yo-Sung Ho, Lossless Data Hiding Based on Histogram Modification of Difference Images, PCM 2004, LNCS 3333 ,2004, 340–347. https://doi.org/10.1007/978-3-540-30543-9_43

[4]. Xuan, G., Zhu, J., Chen, J., Shi, Y., Ni, Z., Su, W.: Distortionless data hiding basedon integer wavelet transform. IEEEElectrionics Letters ,2002, 1646–1648.https://doi.org/10.1049/el:20021131

[5]. GuorongXuan, Qiuming Yao, ChengyunYang, JianjiongGao, Peiqi Chai, Yun Q. Shi, Zhicheng Ni, Lossless Data Hidding Using Histogram Shifting Method Based on Integer Wavelets,Proc. 5th Digital watermarking workshop, IWDW 2006, Korea, vol. 4283, pp. 323-332.https://doi.org/10.1007/11922841_26

[6]. Ho Thi Huong Thom, Ho Van Canh, Trinh Nhat Tien, Steganalysis to Reversible Data Hiding, Proceedings of FGIT 2009 (the Future Generation Information Technology Conference) on Database Theory and Application, Springer-Verlag, Jeju Island, Korea (2009), pp. 1- 6. https://doi.org/10.1007/978-3-642-10583-8_1

[7]. F.-J. Huang and Y. LeCun, Large-scale learning with svm and convolutional nets for generic object categorization, in Proc. Computer Vision and Pattern Recognition Conference (CVPR’06). IEEE Press, 2006. https://doi.org/10.1109/cvpr.2006.164

[8]. D. C. Ciresan, U. Meier, J. Masci, and J. Schmidhuber, A committee of neural networks for traffic sign classification,” in International Joint Conference on Neural Networks, to appear, 2011. https://doi.org/10.1109/ijcnn.2011.6033458

[9]. Stamatios Lefkimmiatis, Non-local Color Image Denoising with Convolutional Neural Networks, in IEEE Conference on Computer Vision and Pattern Recognition 2017 (CVPR 2017). https://doi.org/10.1109/cvpr.2017.623

[10]. Y.A. Alotaibi, High performance Arabic digits recognizer using neural networks, in Proceedings of the International Joint Conference on Neural Networks, 2003. https://doi.org/10.1109/ijcnn.2003.1223444

[11]. H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations, in Proceedings of the 26th International Conference on Machine Learning, 2009, pp.609–616. https://doi.org/10.1145/1553374.1553453

[12]. M. D. Zeiler, D. Krishnan, G. W. Taylor, and R. Fergus, Deconvolutional Networks, in Proc. Computer Vision and Pattern Recognition Conference (CVPR 2010), 2010. https://doi.org/10.1109/cvpr.2010.5539957

[13]. Hyeonseob Nam, Bohyung Han, Learning Multi-domain Convolutional Neural Networks for Visual Tracking, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR2016), https://doi.org/10.1109/cvpr.2016.465

[14]. D. C. Ciresan, U. Meier, L. M. Gambardella, and J. Schmidhuber. Convolutional neural network committees for handwritten character classification. In International Conference on Document Analysis and Recognition (2011), pages 1250–1254. https://doi.org/10.1109/icdar.2011.229

[15]. CBIR image database, University of Washington, available at: http://www.cs.washington.edu/research/imagedatabase/groundtruth/

[16]. USC-SIPI Image Database, http://sipi.usc.edu/services/database/Database.html.

Author'sBiography

Ho Thi Huong Thom: Ho Thi Huong Thom received the B.S. degree of Information Technology department from Haiphong Private University, the M.S. degree and PhD degree in Information Systems from College of Technology, Vietnam National University in Vietnam, in 2001, 2005 and 2012, respectively. She started her career as Lecturer in Department of Information Technology in Haiphong Private University, Vietnam and served 12 years. From 2014, She has been teaching at Vietnam maritime university (VMU).

Nguyen Kim Anh: Nguyen Kim Anh obtained her master's degree in 2009 in Computer Science and Engineering from Centre for Development of Advanced Computing, Noida, India and her bachelor's degree in 2005 in Information Technology form Vietnam Maritime University. She is currently working as a lecture in Information Technology Department in Vietnam Maritime University.

Bui Dinh Vu: Bui Dinh Vu obtained a bachelor's degree in Information Technology from Hanoi National University in 1998. He obtained his master's degree in Computer Science from Hanoi Military Technical Academy in 2006. He currently working at the Faculty of Information Technology, Vietnam Maritime University.

How to Cite

Thom, Ho Thi Huong, Anh, Nguyen Kim and Vu,Bui Dinh "Steganalysis for Reversible Data Hiding Based on Neural Networks and Convolutional Neural Networks",*International Journal of Machine Learning and Networked Collaborative Engineering,* Vol. 02 No. 02, 2018,pp.40-48. doi: https://doi.org/10.30991/IJMLNCE.2018v02i02.001.

Vol. 02, No. 2,(2018), 049–057 , **ISSN: 2581-3242**

IJMLNCE JOURNAL

International Journal of Machine Learning and Networked Collaborative Engineering

Journal Homepage: http://www.mlnce.net/home/index.html

DOI : https://doi.org/10.30991/IJMLNCE.2018v02i02.002

Prediction Model for Pollutants with Onboard Diagnostic Sensors in Vehicles

[a]Bryce Maldonado[*], [b]Malik Bennabi, [c]Vicente García-Díaz, [d]Cristian González García and [e]Edward Rolando Núñez-Valdez

[a]University of South Florida, Florida, USA

maldonado@mail.usf.edu

[b]University of Sakarya, Faculty of Computer and Information Sciences, Sakarya, Turkey

malik.bennabi@ogr.sakarya.edu.tr

[c,d,e]Department of Computer Science, University of Oviedo, Spain

[c]garciavicente@uniovi.es, [d]gonzalezcristian@uniovi.es, [e]nunezedward@uniovi.es

Abstract

In this work, a prediction model is developed to illustrate the relationship between the internal parameters of a vehicle and its emissions. Vehicles emit various hazardous pollutants and understanding the influence of in-vehicle parameters is key to reducing their environmental impact. The values of the internal parameters were collected through the On-Board Diagnostics port, while the values of the emissions were measured from the exhaust pipe using Arduino sensors. The observed values were then matched based on the timestamps received from both sources and fit with both linear and polynomial regressions to accurately model the relationship between the internal parameters and pollutants. These models can then be used to estimate vehicle emissions based on the in-vehicle parameters, including vehicle speed, relative throttle position, and engine revolutions per minute. A wide majority of the relationships between various in-vehicle parameters and emissions show no observable correlation. There are observable correlations between carbon dioxide emissions and vehicle speed, as well as carbon dioxide emissions and engine revolutions per minute. These relationships were modelled using linear and polynomial regression with a resulting adjusted R-squared value of approximately 0.1.

Keywords

41A05,

41A10,

65D05,

65D17,

Artificial Intelligent,

Prediction model,

Sensors,

OBD

1. Introduction

The concentration of car emissions is a crucial factor for manufacturers to consider when designing and manufacturing vehicles. The burning of fossil fuels to power these automobiles creates various

[*] Corresponding author

Bryce Maldonado

Email: maldonado@mail.usf.edu

emissions, some of which are toxic and hazardous to human health [1,2]. Standards and regulations have been developed and refined over the past two decades to address various health and environmental concerns, leading to the creation of institutions and agencies such as the Environmental Protection Agency (EPA) in the United States and the European Environment Agency (EEA) in Europe, to ensure that manufacturers adhere to certain protocols.

Reducing emissions goes beyond the manufacturer and the composition of the fuel used [3], as consumers can participate in eco-driving to further reduce emissions. Various eco-driving programs provide tips and guidelines on methods and techniques to further reduce vehicle emissions by 10% on average [4]. These eco-driving programs emphasize the importance of anticipating traffic, removing unneeded heavy loads from the vehicle, and other small behavioral changes that can reduce emissions [5].

There are many driving factors that influence the fuel consumption of a vehicle, and therefore the emissions, but the plausibility and viability of these eco-driving options may vary between individuals. Eco-driving strategies are being tested and implemented to assist drivers in reducing their fuel consumption, especially in poor driving conditions including severely congested roadways [6]. To aid consumers desiring to participate in eco-driving, it is important to understand the correlations between various in-vehicle parameters and the emissions of a vehicle. Some internal parameters are affected by driving patterns and characteristics, and thus understanding the influence these parameters have on emissions can lead to new eco-driving strategies to be developed. The accessibility and ease of eco-driving can always be improved upon and attract and encourage drivers to reduce their fuel consumption in the long-run [5].

In this paper, various in-vehicle parameters of a car are observed and plotted against the outputted emissions from the exhaust pipe to determine the effect these parameters have on vehicle emissions. Any possible correlations between in-vehicle parameters and the concentrations of outputted emissions are analyzed through simple regressions [7]. This results in prediction models that can be used to determine the influential strength of various in-vehicle parameters on vehicle emissions.

The rest of this work is structured as follows: Section II examines the common pollutants in vehicle emissions and explores previous studies that have examined the significance of various vehicle and driving characteristics on emissions and fuel consumption. Section III describes the method and the reasoning behind the method for collecting the data. The implementation of the methodology is described afterwards. Section IV details the results and the analysis conducted on the recorded data. Section V presents the conclusions drawn from the study and Section VI explores ways and directions to further this work.

2. Background

Driving styles can be classified and deconstructed into identifiable characteristics to establish relationships between driving patterns and fuel consumption [8]. Internal car factors, such as acceleration, deceleration, average speed, RPM, and throttle position have been correlated directly to fuel consumption [8,9]. Managing steady non-volatile speeds and adhering to certain velocity ranges has yielded up to a 20% decrease in vehicle emissions in simulation, with significant reduction still observed in experimental cases 10. Sixty-two observable driving pattern parameters have been narrowed down to nine that have a significant effect on emissions 8. The parameters that have significant influence on emissions can be summarized by five key components: acceleration with high power demand, speed oscillation, extreme acceleration, deceleration, and stop 8. These components were used, in part, for determining which internal vehicle parameters to observe and plot with the emission concentrations when searching for relationships in the data because previous studies examine the effects of driving patterns on fuel consumption instead of emission concentrations directly.

There's been a demonstrated linear positive correlation between RPM and fuel consumption, and a polynomial positive correlation between relative throttle position and fuel consumption using statistical data analysis 9. This indirectly implies a relationship between these internal factors and emission concentrations through fuel consumption. The goal of this study is to expand upon this work to examine more than these two internal vehicle parameters, and to note their direct impact and emission concentrations. Past studies have examined the accuracy of On-Board Diagnostic systems using the observable concentrations of emissions from the exhaust pipe 11. Rather than comparing the observed concentrations of emissions to specific internal vehicle parameters, these outputted concentrations were classified by driving states (accelerating, cruise, idle) 11. This paper further seeks to examine the direct relationship between chosen internal vehicle factors and emission concentrations, modeling any observed correlations between these factors and emission concentrations to accurately estimate emission levels when driving.

Vehicle emissions mainly consist of nitrogen, carbon-dioxide, and water vapor 12; hazardous emissions such as carbon monoxide, nitrogen-oxide compounds, hydrocarbons, and particulate matter constitute less than 1% of the total emissions from motor vehicles 12. Most of the sensors for this study will examine the hazardous emissions, but a carbon dioxide sensor will also be utilized to examine a larger portion of emissions. A tool has been developed for simulating the concentrations of these emissions given various other internal parameters 13. The simulator uses a verified database of emissions concentrations to simulate emissions, instead of using real-time observed data to correlate the internal parameters to emission concentrations 13. A statistical model for vehicle emissions has been created previously 14, but not using only internal vehicle parameters and real-time observed data.

3. Experiment Description

The On-Board Diagnostics system was used to record the values of various internal car factors while driving. It provides data for many of the current conditions and parameters of a vehicle 15. The On-Board Diagnostics PIDs (OBD-II PIDs) were read using an OBD-port scanner while the vehicle was in use; Table 2 displays the recorded in-vehicle parameters obtained through the OBD-II port. Not all parameters are relevant, while the bolded parameters significantly related to fuel consumption and as such were the chosen to be analyzed. These parameters were chosen according to the key components of strong fuel consumption influence 8. To measure the emission concentrations, a set of MQ gas sensors were connected to an Arduino board, as seen in Figure 1; Table 2 provides the names of the sensors employed in the experiment and the gases they detected and measured. Automobiles produce a variety of emissions including carbon dioxide (CO2), hydrocarbons (HC), nitrogen-oxide compounds (NOx), and other particulate matter 12. Some MQ sensors used in the experiment detected gases that were not present or strongly represented, in standard vehicle emissions, such as alcohol; hydrocarbons and carbon monoxide make up less than 1% of the total emissions of a motor vehicle, while carbon dioxide constitutes a more observable 12% 12. Because of its observability, the MG811 sensor is the most significant sensor to monitor; the other hydrocarbon and carbon monoxide sensors were included to note any significant relationships between driving parameters and the more toxic vehicular emissions.

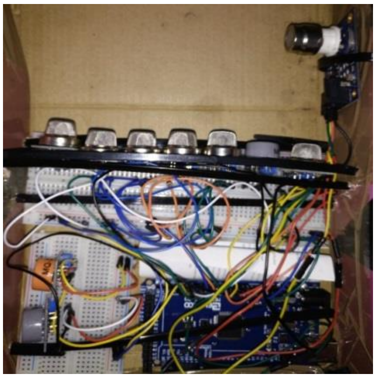


Fig.1. Arduino MQ2-9,135 & MG811 Sensors

Fig.2. Citroen C4 during data collection

4. Implementation

Three different automobiles were driven during the experiment to examine the influence in-vehicle parameters have on emissions regardless of the manufacturer and model of car; the chosen automobiles included a Citroen C4, a Citroen Xsara, and a Honda Accord. The vehicles tested in this experiment ran on diesel and gasoline. During the trial run with each vehicle, an Arduino wired with the set of MQ gas sensors was secured into a cardboard holding device and fastened to the end of the exhaust pipe, as seen in Figure 2. The automobile was driven in a nondiscriminatory manner around the city of Oviedo, Spain for an undetermined and variable length of time. Data recorded from both the OBD-II port and the MQ-sensors was collected from each vehicle and then synchronized through the timestamps included in the data. There is no fixing of Arduino values to account for possible delay between a change in an internal parameter and a change in the emissions out of the exhaust pipe. To justify this, a second pass between each full reading of the MQ sensors on the Arduino.

Table 1. Observed vehicle parameters

Internal Parameter	Units
Accelerometer(X)	m/s^2
Accelerometer(Y)	m/s^2
Accelerometer(Z)	m/s^2
Accelerometer(Total)	m/s^2
GPS Speed	m/s
OBD Speed	m/s
Intake Pressure(MAP)	Pa
RPM	rpm
Throttle Position	%
Relative Throttle Position	%

Table 2. Tested Sensors

Sensor	Gas(es) detected
mq2	Combustible gas, smoke
mq3	Alcohol
mq4	Methane, Propane, Butane
mq5	Butane, Propane, Methane
mq6	Liquefied petroleum, Butane, Propane, LPG
mq7	Carbon monoxide
mq8	Hydrogen
mq9	Carbon Monoxide, Methane
mq135	Ammonia sulfide, Benzene vapor
mg811	Carbon dioxide

5. Results

An overwhelming majority of the graphs displayed no significant observable relationship or correlation between an internal parameter and any of the MQ gas sensors. Some examples of these graphs are shown in Figures 3 and 4. All the parameters from Table 2 were examined, but only a few relationships produced graphs that demonstrated a perceivable correlation. These correlations can be seen in Figure 5 and 6 which clearly demonstrate a positive correlation between the independent and dependent variables, but the large variance of the data, especially at higher values, lessens the feasibility of an accurate estimation model. Regardless, a simple linear (shown in red) and polynomial regression (shown in blue) are run on each graph to model the general trend of the data. The resulting prediction model equations are given in Table 3, along with the adjusted R-squared value (ARV) and the residual standard error (RSE). The wide variance noted above in the plots has resulted in none of the adjusted R-squared values reaching over 0.2, meaning that the prediction models have an approximate accuracy of 17-20%. For the other emissions that the remaining sensors tested for, including CO, HC, and combustible smoke, there appeared to be no significant correlation between internal vehicle parameters and those emissions in this experiment.

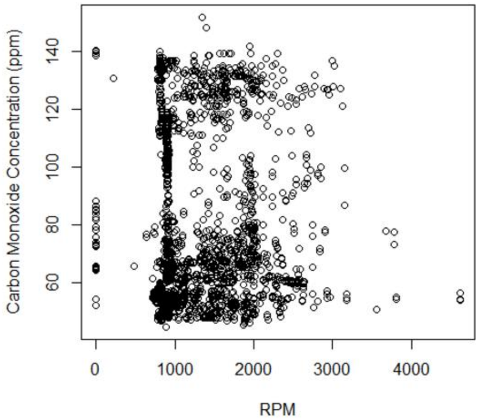


Fig. 3. RPM vs Carbon Monoxide Concentration

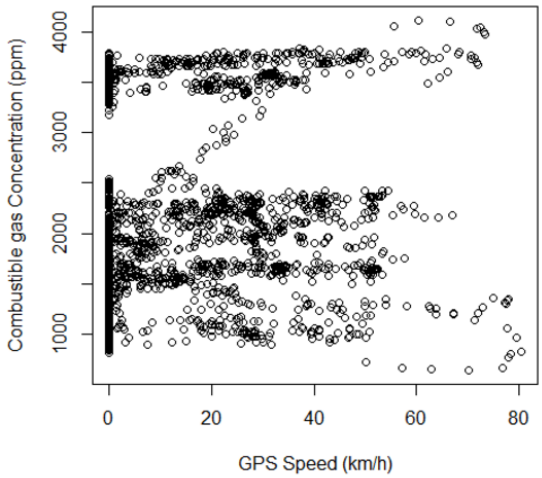


Fig. 4. GPS Speed vs Carbon Monoxide Concentration

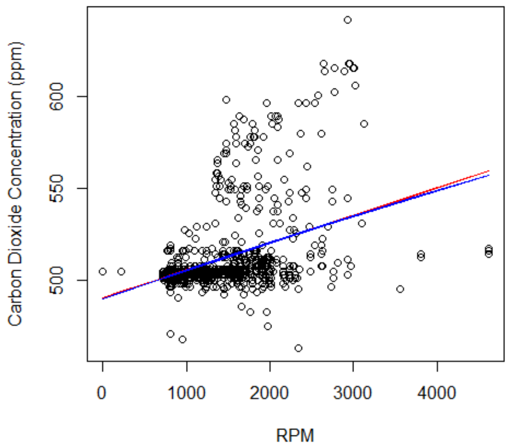


Fig. 5. RPM vs Carbon Dioxide Concentration

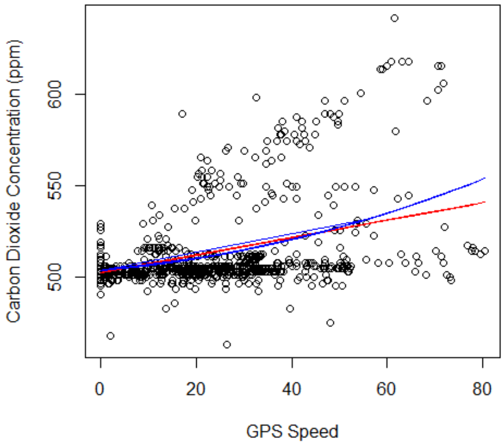


Fig. 6. GPS Speed vs Carbon Dioxide Concentration

6. Discussion

The lack of significant and accurate estimation results may stem from the method by which the data was collected. Having MQ gas sensors attached to the exhaust pipe of the vehicle may have led to volatile data collection since, as the vehicle moves, the emissions may move past the sensors before the MQ sensors have time to detect them. A total of 10 Arduino gas sensors were used during the experiment, and each sensor took 0.1 seconds to produce a new piece of data, meaning each sensor only took a reading from the exhaust pipe every second. This means that a vehicle moving at just 10 km/h would travel 2.77 meters before each sensor would read from the exhaust pipe. Additionally, while the Arduino sensors were secured on the opening of the exhaust pipe, it is possible for emissions to miss the sensors due to the nature of gases and the wind at high velocities. Gases expand to fit the size of their container, emissions may have expanded into the atmosphere once outside of the exhaust pipe and missed the MQ sensors.

Table 3. Regression Analysis

Regression type	Equation	ARV*	RSE#
Linear Regression RPM vs CO2	490.9+0.01484x	0.1762	18.11
Polynomial Regression RPM vs CO2	$490.3+0.01574x-2.663e\text{-}07x^2$	0.1756	18.12
Linear Regression GPS Speed vs CO2	502.788+0.47524x	0.1854	18.01
Polynomial Regression GPS Speed vs CO2	$504+0.2182x+0.00501x^2$	0.1942	17.92

*ARV–Adjusted R-squared Value
#RSE–False Residual Standard Error

7. Conclusions and Future work

Understanding the relationship between driving patterns and vehicle emissions is important when furthering eco-driving efforts. An estimation model for predicting and illustrating the relationships between internal vehicle parameters and emissions was sought in this experiment, to observe the significance of these parameters€™ influence on emissions. The data collected from the OBD-II port and the MQ sensors yielded significant relationships only between a couple internal parameters and the carbon dioxide emissions.

Hydrocarbons, carbon monoxide, and combustible smoke emissions were not significantly affected by variations in the recorded internal vehicle parameters.

There is a general positive correlation between engine RPM and carbon dioxide emissions, as well as vehicle speed and carbon dioxide emissions. This positive correlation has been fitted with estimation models from both linear and polynomial regressions. The resulting estimation models have adjusted R-squared values that are less than 0.2. While these estimation models do not accurately predict emission values, they do illustrate the positive relationship between the two variables, which is significant enough to be observed when plotted.

There exist multiple possible directions to continue the work of this experiment. Nitrogen oxide emissions constitute the largest proportion of vehicle emissions, this experiment could be repeated with nitrogen oxide and other hydrocarbon sensors to continue to search for significant relationships. Increasing the read time of the sensors and trapping emissions to collect data before releasing them into the air could provide more desirable results.

More tests can be run including a wider variety of vehicle makes and models. The vehicle selection can also be expanded by including hybrid vehicle, such as diesel-electric, and analyzing emissions with respect to in-vehicle parameters and fuel type.

8. Acknowledgements

This research has been partially supported by the National Science Foundation under grants No. 1458928 and No. 1645025, An REU Site on Ubiquitous Sensing.

The authors would like to thank the University of Oviedo, the University of South Florida, USA for organizing the research camp and providing us with all the support and assistance needed along this work.

The University of Sakarya of Turkey, for the indirect support of the participation of Malik Bennabi in this work.

All the professors who participated in the elaboration of the research camp content, activities, workshops and mentoring as well, particularly Dr. Miguel A. Labrador & Dr. Daniel F. Lanvin.

References

[1].Pearson RL, Wachtel H, Ebi KL. Distance-weighted traffic density in proximity to a home is a risk factor for leukemia and other childhood cancers. *J Air Waste Manage Assoc*. 2000;50(2):175-180. doi:https://doi.org/10.1080/10473289.2000.10463998.

[2]. Nicolai T, Carr D, Weiland SK, et al. Urban traffic and pollutant exposure related to respiratory outcomes and atopy in a large sample of children. *Eur Respir J*. 2003;21(6):956-963. doi:https://doi.org/10.1183/09031936.03.00041103a.

[3]. Zielinska B, Sagebiel J, McDonald JD, Whitney K, Lawson DR. Emission rates and comparative chemical composition from selected in-use diesel and gasoline-fueled vehicles. *J Air Waste Manage Assoc*. 2004;54(9):1138-1150. doi:https://doi.org/10.1080/10473289.2004.10470973.

[4]. Barkenbus JN. Eco-driving: An overlooked climate change initiative. *Energy Policy*. 2010;38(2):762-769. doi:https://doi.org/10.1016/j.enpol.2009.10.021.

[5]. Beusen B, Broekx S, Denys T, et al. Using on-board logging devices to study the longer-term impact of an eco-driving course. *Transp Res part D Transp Environ*. 2009;14(7):514-520. doi:https://doi.org/10.1016/j.trd.2009.05.009.

[6]. Saboohi Y, Farzaneh H. Model for developing an eco-driving strategy of a passenger vehicle based on the least fuel consumption. Appl Energy. 2009;86(10):1925-1932. doi:https://doi.org/10.1016/j.apenergy.2008.12.017.

[7]. Friedman J, Hastie T, Tibshirani R. *The Elements of Statistical Learning*. Vol 1. Springer series in statistics New York; 2001.

[8]. Ericsson E. Independent driving pattern factors and their influence on fuel-use and exhaust emission factors. *Transp Res Part D Transp Environ*. 2001;6(5):325-345. doi:https://doi.org/10.1016/S1361-9209(01)00003-7.

[9]. Lee MG, Park YK, Jung KK, Yoo JJ. Estimation of fuel consumption using in-vehicle parameters. *Int J u-and e-Service, Sci Technol*. 2011;4(4):37-46.

[10]. Barth M, Boriboonsomsin K. Energy and emissions impacts of a freeway-based dynamic eco-driving system. *Transp Res Part D Transp Environ*. 2009;14(6):400-410. doi:https://doi.org/10.1016/j.trd.2009.01.004.

[11]. Frey HC, Unal A, Rouphail NM, Colyar JD. On-road measurement of vehicle tailpipe emissions using a portable instrument. *J Air Waste Manage Assoc*. 2003;53(8):992-1002. doi:https://doi.org/10.1080/10473289.2003.10466245.

[12]. Reşitoğlu, İbrahim Aslan KAK. The pollutant emissions from diesel-engine vehicles and exhaust aftertreatment systems. *Clean Technol Environ Policy*. 2015;17(1):15-27. doi:https://doi.org/10.1007/s10098-014-0793-9.

[13]. Smit R. Development and performance of a new vehicle emissions and fuel consumption software (PΔP) with a high resolution in time and space. *Atmos Pollut Res*. 2013;4(3):336-345. doi:https://doi.org/10.5094/APR.2013.038.

[14]. Cappiello A, Chabini I, Nam EK, Lue A, Zeid MA. A statistical model of vehicle emissions and fuel consumption. In: *Intelligent Transportation Systems, 2002. Proceedings. The IEEE 5th International Conference on*. ; 2002:801-809.

[15]. Cueva-Fernandez G, Espada JP, García-Díaz V, García CG, Garcia-Fernandez N. Vitruvius: An expert system for vehicle sensor tracking and managing application generation. *J Netw Comput Appl*. 2014;42(1):178-188. doi:http://dx.doi.org/10.1016/j.jnca.2014.02.013.

Author's Biography

Bryce Maldonado: Bryce Maldonado is a teaching assistant in the Department of Computer Science and Engineering at the University of South Florida (USA). He is a graduate student with a BS in Computer Science and Engineering having graduated from the College of Engineering at the University of South Florida in 2017.

Malik Bennabi: Malik Bennabi is a senior student of Information Systems Engineering at the University of Sakarya (Turkey). He has been a student of Electrical and Electronic Engineering at the IEEE University of Boumerdès (Algeria). He has completed the Cisco CCNA in 2011. He has been an Erasmus+ scholar at the School of Computer Engineering, University of Oviedo (Spain) during 2016-2017. He has successfully completed several Entrepreneurship programs in different countries. He is a co-founder of the startup I-sen in Asturias (Spain). He is a winner of several entrepreneurship and innovation competitions around the world and has been a mentor at several hackathons including the IoT Olympiads Boston 2014 (USA). He is a polyglot speaking fluently 7 languages, comfortable with 2 and continually learning. He successfully completed the summer research program provided by the University of South Florida (USA) and the University of Oviedo (Spain), summer 2017. His research interests include Machine and Deep learning, Artificial intelligence, Blockchain, Entrepreneurship and Innovative learning.

Vicente García Díaz: Vicente García Díaz is an associate professor in the Department of Computer Science of the University of Oviedo (Spain). He has a PhD from the University of Oviedo in Computer Engineering. He has been a visiting professor in the University of Manchester (UK), Distrital University Francisco José de Caldas (Colombia) and the University of Lisbon (Portugal) and taught courses on computer science in other international institutions. He has published more than 70 research papers in international conferences and journals. He has graduated in prevention of occupational Risks and is a Certified associate in project management through the Project Management Institute. His research interests include domain specific languages, data science, user of computers for learning and teaching, machine learning and extraction of knowledge from data, extracted from social networks or internet of things-based networks.

Cristian González García: Cristian González García is an assistant professor in the Department of Computer Science of the University of Oviedo (Spain). He is a Technical Engineer in Computer Systems, M.Sc. in Web Engineering, and a Ph.D. in Computers Science graduated from School of Computer Engineering of Oviedo in 2011, 2013, and 2017 (University of Oviedo, Spain). He has been a visitor Ph.D. candidate in the University of Manchester, United Kingdom. Besides, he has been working in different national and regional projects, in projects with private companies, and in the University of Oviedo as a professor. He has published 10 journals articles, 4 conference articles, and 2 book chapters.

Edward Rolando Núñez Valdez: Edward Rolando Núñez Valdez is an assistant professor in the Department of Computer Science of the University of Oviedo (Spain). He is Ph.D. from the University of Oviedo in Computer Engineering. Master in Software Engineering from the Pontifical University of Salamanca and B.S. in Computer Science from Autonomous University of Santo Domingo. He has participated in several research projects, He has taught computer science at various schools and universities and has worked in software development companies and IT Consulting for many years. He has published several articles in international journals and conferences. His research interests include Web Engineering, Artificial Intelligence, Recommendation Systems, Modeling Software with DSL and MDE.

How to Cite

Maldonado, Bryce, Bennabi, Malik, García-Díaz, Vicente, García, Cristian González and Valdez, Edward Rolando Núñez-, "Prediction Model for Pollutants with Onboard Diagnostic Sensors in Vehicles", *International Journal of Machine Learning and Networked Collaborative Engineering,* Vol. 02 No. 2, 2018, pp. 49-57. doi : https://doi.org/10.30991/IJMLNCE.2018v02i02.002.

Vol. 02, No. 2 ,(2018), 058–066 , **ISSN:2581-3242**

IJMLNCE JOURNAL

International Journal of Machine Learning and Networked Collaborative Engineering

Journal Homepage: http://www.mlnce.net/home/index.html

DOI : https://doi.org/10.30991/IJMLNCE.2018v02i02.003

Information Processing in Neuron with Exponential Distributed Delay

[a]Saket Kumar Choudhary[*], [b]Sunil Kumar Bharti

[a]*Impel Labs Pvt. Ltd., Bengaluru – 560102, saket_au@rediffmail.com; saket@impel.io, ORCID: 0000-0002-0448-3076*

[b]Central University of Haryana, Mahendragarh – 123031; sirbharti@gmail.com *ORCID: 0000-0002-0245-6496*

Abstract

Artificial intelligence (AI) has been become the primary need in nearly all sectors namely engineering, services, banking, finance, defense, space etc [3], [33]. Artificial intelligence in these sectors can be implemented in two ways: (i) hardware level implementation (ii) software level implementation. Both kinds of AI implementation require neuron models which mimic the minimal set of real neuron functionality. To this end, Leaky Integrate-and-Fire (LIF) model is performing as the backbone for both kinds of AI implementation. At hardware level implementation, it's a variant, called as neuristors, is used at chip level implementation, whereas a number of variants LIF model are used to implement AI at software level. In this work, the extended LIF model in distributed delay kernel regime is analyzed. The impact of exponentially distributed delay (EDD) memory kernel on spiking activity and steady state membrane potential distribution (SVD) of LIF neuron is investigated. Fokker-Planck equation associated with the considered model is solved to investigate SVD of the neuron in sub-threshold regime, which results Gaussian distribution. In order to study the information processing, spiking activity of the model is investigated, which is further extended to neuronal rate-code scheme. These finding have been compared with simple LIF model with stochastic input. It is evident that steady state membrane potential distribution of the LIF neuron is invariant due to the presence of EDD. Such kinds of neuron models are useful to implement artificial neural networks. To this end, the proposed model can used to implement recurrent neural networks (RNN) with comparatively more accuracy. Similarly, this model can also be investigated in term of chip level implementation of AI.

Keywords

Artificial Intelligence,

Distributed Delay,

Fokker-Planck Equation,

LIF Model,

LSTM,

Spiking Activity,

Recurrent Neural Network,

Steady State Probability Distribution

1. Introduction

Artificial intelligence is computer based system comprised of neural networks, which behaves like human brain [29]. Neural networks are artificial human-brain like computing structure which incorporates

[*] Corresponding author

Saket Kumar Choudhary

Email: saket_au@rediffmail.com; saket@impel.io

the membrane information processing mechanism [7], [9]. Long short-term memory (LSTM) network, recurrent neural network are two important neural networks which deals with the memory element. Neuron forms basis for information processing in these networks and generates the real central nervous system information processing behavior [1] [7]. A biological neuron has three fundamental information processing entities, namely, soma, axons and dendrites [7] [28]. Dendrite and axon meet at synapse to establish a communication link between two neurons [1], [28]. Neuron generates action potential in form spike sequences which travel through axon down the nervous system [4], [15]. In this way, a neuron forms a computational device which maps synaptic input into spikes sequence. This has motivated to researchers to study the correspondence among input stimuli and spiking patterns. Computational neuroscience provides a platform to understand brain information processing functions in terms of representation, information processing as chemical and electrical signals, effect of memory element [26] so that artificial implementation become more robust and accurate. Information theoretic framework includes mechanisms to measure uncertainty between encoding and decoding of spiking patterns in quantitative means [4], [7], [13].

Lapicque (1907) has proposed the integrate-and-fire neuron model (IF model), first neuron model, equivalent to an electrical RC-circuit with an additional threshold constraint, i.e. neuron potential resets a lower value as soon as it reaches to a certain value (threshold) [1], [7], [9]. Under the influence of noisy input stimulus and noisy environment, neuron reflects the highest level of variation in spiking pattern [25]. Sub-threshold and super-threshold regime are two set of parameters in neuronal information processing which covers the noise. Neuron remains silent in former regime whereas it emits spikes in the late regime under no influence of noise [17].

Memory elements can be introduced in a neuron model via kernel function. Karmeshu et. al. [20] has proposed a distributed delay framework and has made an enlarged study of LIF model for exponential distributed delay kernel function and noticed oscillatory spiking patterns in inter-spike-interval distribution [20]. In this article, LIF model with exponentially distributed kernel function is further investigated in terms of spiking activity and SVD. Simulation based study is done to study the spiking patterns and SVD is computed by solving the associated Fokker-Planck equation (FPE). These findings are compared with stochastic LIF model.

The article is organized into six sections. After a brief introduction of neural network, neuron model and its information processing mechanism in Section I, neuronal DDF is explained in Section II. Section III deals with computation of SVD. Spiking activity is analyzed via neuronal information processing rate-code scheme in Section IV. Section V discusses about artificial intelligence in terms of various used techniques and its application. Finally, Section VI contains analysis of findings and conclusion.

2. Distributed Delay Framework Of Neuron Model

Membrane potential is the key parameter for neuron models. Variation in membrane potential is always taken care to deal with neuron’s behavior. LIF model is an extended version of IF model. It assume that the membrane conductance is leaky i.e. if there is no electrical activity happing to and fro to the neuron, then membrane potential will get decayed with respect to time. To and for activity of membrane potential in a neuron occur due to the movement of various ions and molecules. The aggregated effect of these ions and molecules can be captured in term of stochastic current. Thus, the rate of change of membrane potential ($V(t)$) for LIF model with stochastic input stimulus is defined as [5], [6], [7], [10].

$$\frac{dV(t)}{dt} = -\beta V(t) + \mu + \xi(t) \tag{1}$$

Here, β is membrane decay constant, μ is mean input stimulus, $\xi(t)$ is delta correlated Gaussian White noise with intensity , i.e. $\langle \xi(t) \rangle = 0$ and $\langle \xi(t_i)\xi(t_j) \rangle = \frac{\sigma^2}{2}$. DDF for the LIF model is defined as [7], [20]

$$\frac{dV(t)}{dt} = -\beta \int_0^t K(t-\tau)V(\tau)d\tau + \mu + \xi(t) \tag{2}$$

Here $K(t)$ is distributed delay kernel. This captures the effect of past membrane potentials on the current values. $K(t)$ can has multiple forms, like Gamma distribution, Sigmoid, Laplace distributed etc. Substitution of gamma distributed delay kernel for $K(t)$ in Eq. (2) results as below [20].

$$\frac{dV(t)}{dt} = -\beta\int_0^t \frac{\eta^{m+1}(t-\tau)^m e^{-\eta(t-\tau)}}{m!} V(\tau)d\tau + \mu + \xi(t) \tag{3}$$

Here η and m are gamma distributed kernel parameters. Eq. (3) reduces into LIF model with exponential distributed delay kernel for $m=0$. Exponential distributed delay kernel is also known as the weak-kernels, as it captures only the first order of membrane potential delay dependency. This results membrane potential evolution process $\{V(t);t\geq 0\}$ into a non-Markovian process. Analytical analysis of a non-Markovian process is a too complex and has to look on some other methods. The present non-Markovian process can be extended into higher dimensions to transform membrane potential evolution into Markovian process [20]. LIF model with EDD in extended space takes the following form [20].

$$\frac{dV}{dt} = -\{\eta\beta U_0(t) - \mu\} + \xi(t) \tag{4.1}$$

$$\frac{dU_0}{dt} = -\eta\{U_0(t) - V(t)\} \tag{4.2}$$

with initial conditions $V(t)=V_0, U_0(t)=0$ at $t=0$.

Eqs. (4.1) and (4.2) form system of coupled stochastic differential equations and can analyzed analytically. Ito method and Stratanovich method are two important techniques to numerically study stochastic differential equations. The primary difference in both techniques is the point of approximation in the sub-time interval. Ito technique uses left bound of sub-time interval whereas Stratanovich techniques uses mean of the left and right bounds of sub-time interval [14].

3. SVD for LIF Neuron with EDD Kernel Function

Rate of change of spatial probability of a stochastic variable are studied in terms of Fokker-Planck equations (FPEs). For the system of coupled stochastic Eqs. (4.1) and (4.2), FPE results membrane potential probability distribution [5], [6], [31]. FPE is also used to solve first passage time problem, firing rate computation and estimation of moments of stochastic conductance's [21], [23]. Solution of Fokker-Planck equation highly depends on the boundary conditions [14]. FPE solution highly depends on initial conditions and boundary conditions. When membrane potential reaches at threshold value, the probability of generation of spike is 1 for absorbing boundary condition, whereas, it is 0 for reflecting boundaries. In later condition, probability current flux $J(V,t)$ becomes equal to zero. The extended LIF model, defined in Eqs. (4.1) and (4.2) are analyzed with reflecting boundary condition.

Let $p^E(V,U_0,t)$ is membrane potential probability distribution in extended apace. Following Burkitt [5], [6] and Frank [14], corresponding Fokker-Planck equation becomes

$$\frac{dp^E}{dt} = \frac{\partial}{\partial V}(\beta U_0 - \mu)p^E + \frac{\partial}{\partial U_0}\eta(U_0 - V(t))p^E + \frac{\sigma^2}{2}\frac{\partial^2 p^E}{\partial V^2} \tag{5}$$

With boundary conditions: $p^E(V,U_0,t\,|\,t=0) = \delta(t-t_0)\delta(U_0(t)-U_0(t_0))\delta(V(t)-V(t_0))$

and $p^E(V,U_0,0) = Vp^E(V,U_0,t\,|\,t=0) = 0$ as $V\to\infty$.

Eq. (5) can be rewritten with differential operator $\nabla = (\frac{\partial}{\partial V}, \frac{\partial}{\partial U_0})$ as

$$\frac{\partial p^E}{\partial t} = \nabla(Ap^E + (\nabla(Bp^E))^T)\text{; with } A = \begin{pmatrix} \beta U_0 - \mu \\ \eta(U_0 - V) \end{pmatrix} \text{and } B = \begin{pmatrix} \frac{\sigma^2}{2}, 0 \\ 0, 0 \end{pmatrix} \quad (6)$$

The probability current flux takes the form

$$J = \nabla(Ap^E + (\nabla(Bp^E))^T) \quad (7)$$

In case of SVD, $\frac{\partial p^E}{\partial t} = 0$, $p^E = p_s^E$ and for reflecting boundaries, Eq. (7) yields

$$Ap^E + (\nabla(Bp^E))^T = 0 \quad (8)$$

Substituting of A, B and using matrix calculation in Eq. (8) yields

$$\begin{pmatrix} (\beta U_0 - \mu) p_s{}^E + \frac{\partial}{\partial V}(\frac{\sigma^2 p^E{}_s}{2}) \\ \eta(U_0 - V) \end{pmatrix} = 0 \quad (9)$$

Its simplification further results

$$\frac{1}{p^E{}_s} \frac{\partial p_s{}^E}{\partial V} = -\frac{\sigma^2}{2}(\beta V - \mu) \quad (10)$$

Integration of Eq. (10) results SVD as

$$p_s{}^E = K_E \exp\{-\frac{\beta}{\sigma^2}\{(V-\mu)^2 - \mu^2\}\} \quad (11)$$

Law of conservation of probabilities can be used to compute the normalization constant (K_E) [12]. Eq. (11) is the stationary state membrane potential distribution of the LIF model in exponentially distributed delay kernel. This membrane potential distribution has the Gaussian distributed form. Similarly, SVD for LIF neuron with stochastic input stimulus is computed below.

Let $p^L(V,t)$ is the spatial probability distribution of membrane potential for LIF model. FPE associated with Eq. (1) is given as

$$\frac{dp^L}{dt} = \frac{\partial}{\partial V}(\beta V - \mu) p^L + \frac{\sigma^2}{2} \frac{\partial^2 p^L}{\partial V^2} \quad (12)$$

With boundary conditions:

$p^L(V,t \mid t=0) = \delta(t-t_0)\delta(V(t)-V(t_0))$ and $p^L(V,0) = Vp^L(V,t \mid t=0) = 0$ as $V \to \infty$.

Probability current flux J for Eq. (12) becomes

$$J = (\beta V - \mu) p^L + \frac{\sigma^2}{2} \frac{\partial p^L}{\partial V} \quad (13)$$

For SVD of neuron, $\frac{\partial p^L}{\partial t}=0$, $p^L = p_s^L$. Simplification of Eq. (13) for reflecting boundaries results

$$\frac{1}{p_s^{\ L}}\frac{\partial p_s^{\ L}}{\partial V}=-\frac{\sigma^2}{2}(\beta V-\mu) \tag{14}$$

Integration of Eq. (14) yields SVD with normalization constant K_L as

$$p_s^{\ L}=K_L\exp\{-\frac{\beta}{\sigma^2}\{(V-\mu)^2-\mu^2\}\} \tag{15}$$

Eq. (15) is the stationary state membrane potential distribution of the LIF model with stochastic input stimulus. This membrane potential distribution is also Gaussian distributed and similar with the SVD of LIF model with exponential distributed delay kernel.

4. Information Processing in a Neuron with Exponential Distributed Delay

Neuron encodes information in term of spikes. Rate code scheme and temporal code scheme are two neuron information encoding scheme [2], [13], [17], [18], [19], [27], [28]. Rate code scheme uses average number of spikes fired by neuron to encode information [15], [16], [17]. Time interval between two consecutive spikes (inter-spike-interval) is used to encode information in temporal coding scheme [17], [25].

Information processing of LIF neuron in EDD is analyzed with rate code scheme. Euler-Maruyama numerical simulation technique for stochastic differential equation [19] is used to simulate neuronal models represented in Eq. (1) and Eq. (4). The average number of spikes per $\sec$ fired by neuron (Spike Count) with different values of parameters μ, σ and η is calculated. In each iteration of simulation, averaging is done for 1000 trials of simulation with $10000\ m\sec$ time in each trial. Findings are illustrated in Fig. 1-Fig. 4.

Parameter value β is taken ranging from 0.01 to 0.3 with step-size of 0.01 [7], [8]. Spike rate for LIF model in EDD is similar to that of LIF model with stochastic input at small β and reduces in LIF model in EDD is slowly than LIF model with stochastic input for increment in β as shown in Fig. 1. Lowering in applied input stimulus μ by half, firing rate for both models reduces, but, the firing activity of LIF model stops for larger values of β whereas, LIF model with EDD remains spiking. This observation is shown depicted in Fig. 2. Similar spiking activity has been observed neuronal disorder like schizophrenia [32].

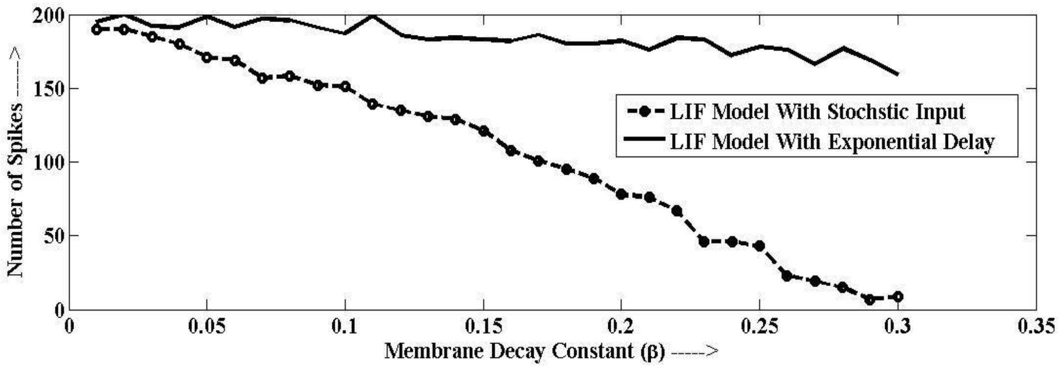


Fig. 1. Spike Count for $\mu=0.2$, $\sigma=0.05$ and $\eta=0.1$

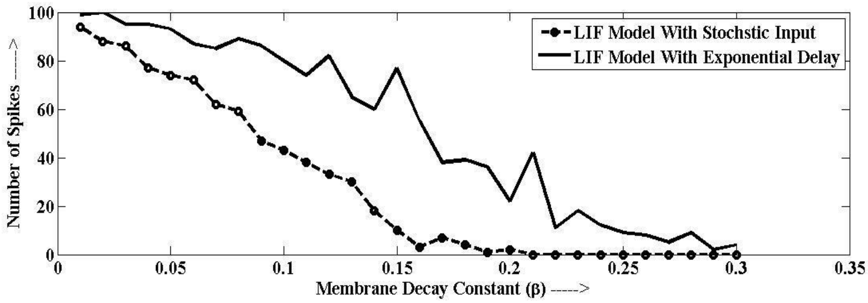


Fig. 2. Spike Count for $\mu = 0.1$, $\sigma = 0.05$ and $\eta = 0.1$

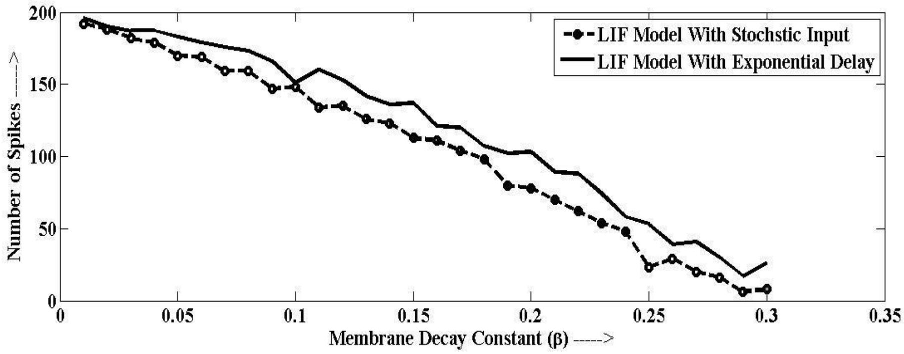


Fig. 3. Spike Count for $\mu = 0.1$, $\sigma = 0.05$ and $\eta = 1$

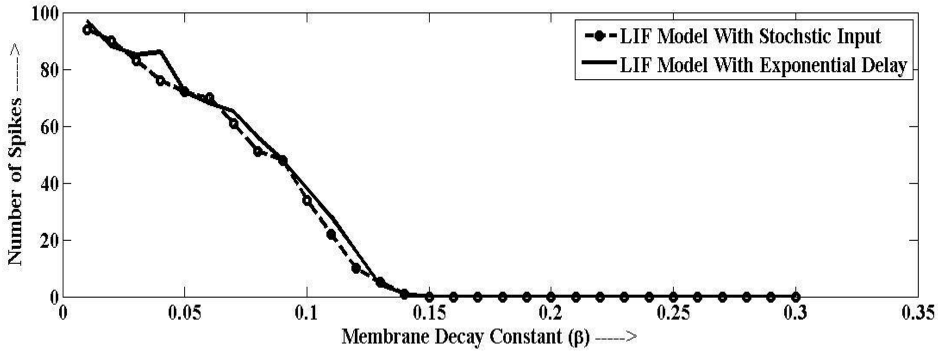


Fig. 4. Spike Count for $\mu = 0.1$, $\sigma = 0.05$ and $\eta = 1.1$

Spiking activity of LIF model with distributed delay decreases with the increase in delay parameter η as shown in Fig. 3 and Fig. 4. For $\eta = 1$, spiking activity of both models result similar patterns. In Fig. 4, it is well illustrated that spiking activity of simple LIF model and LIF model with exponential distributed delay approximately equal which provides the critical values for parameters η at which effect of delay kernel function on evolution of membrane potential becomes negligible.

5. Discussion

Artificial intelligence is a way to transform machines as intelligent machines. This can be done at hardware as well as software level. AI has improved human life in terms of manufacturing automation and service performances in past 20 years [29]. Intelligent machines and systems are also known as expert systems. Expert systems are widely used in business, engineering, medicine, science, security, space, weather forecasting etc. to solve complex problems [29], [33], [11]. McCulloch and Pitts have used artificial neurons to study the AI with logical modeling, first time, in 1965 [29]. Since then, multiple techniques have been evolved to incorporate AI in systems and machines. Few important AI techniques are expert systems, fuzzy logic, artificial neural networks (ANN), multi-layered detectors and learning algorithm and genetic algorithm [29]. Expert system uses person's area of expertise and optimizes the use of their skills. It is highly applicable in biomedical, logistics, marketing, medicines and retails fields. Fuzzy logic and corresponding neural networks uses back propagation and superiorities in combination to incorporate learning and decision-making. It is widely used in time-sequence calculations. Artificial neural network mimics real neuron comparatively more similar way. First of all these networks learns from past data which is termed as training, which includes network structures with different types of connection among multiple artificial neurons. During training process, these connections obtain a weight (synaptic strength) which is used in later stage decision makings. Multiple-layered detectors and learning algorithms involves ANN with more than two artificial neuron layers. Such kind of networks are also termed as deep learning networks and are being used to obtain decision making in multi-class classification problems. Genetic algorithm uses natural selection mechanism and is a probability based AI technique. Genetic algorithms are used to complex problem in economic, information system, mechanical learning, social systems etc. Few important area of AI application is given as below.

- Computer Games: AI helps to artificial character in finding his movement path, making a decision for action and learning from previous movements.
- Banking and Finance Sector: AI helps to find fraud-detection and generating multiple recommendations.
- Medicine: AI helps to clinical support system, medical image classification, diagnostic analysis, MRI brain tumor detection etc.
- Security: AI helps in security in many folds. It is easier to monitor and provide security for human and places with the combination of AI and computer vision. AI is also useful in network monitoring and security in term of intrusion detection and prevention.

6. Conclusion and Future Scope

SVD of LIF neuron in EDD (Eq.(11)) is Gaussian distribution and is asymptotically similar to SVD of LIF neuron with stochastic input (Eq.(15)) i.e. limit $V \to \infty, p_s{}^E \simeq p_s{}^L$. This finding suggests that the exponentially distributed delay kernel function has no effect on stationary state membrane potential of LIF neuron in it's sub-threshold regime. LIF model with stochastic input stimulus and it's extended form in distributed delay regime, both, represents similar spiking activity. Simulation based analysis results that the spike count of neuron increases for EDD. EDD behaves as a memory element and helps neuron to reach at threshold value in quicker time. Masas [24] has suggested a mechanism for self organization and learning from noise in neural networks. EDD has virtue to include memory element which can further be extended to model neural networks. Lim et. al. [22] has suggested a mechanism for hardware level implementation of artificial neural network (ANN) which can further be experimented in EDD.

References

[1]. Abbott, L. F. and Dayan, P., (2001), "Theoretical Neuroscience: Computational and mathematical modeling of neural systems", The MIT press, [ISBN-10: 9780262541855].

[2]. Arcas, y A AB; Fairhall, A. L. and Bialek, W., (2003) "Computation in a single neuron: Hodgkin and Huxley Revisited," Neural Computation, Vol. 15, pp. 1715-1749 [DOI: 10.1162/08997660360675017].

[3]. Belgrano, A., Malmgren, B. A. and Lindahl, O., (2001), "Application of artificial neural networks (ANN) to primary production time-series data", Journal of Plankton Research, Vol. 23, Issue: 6, pp. 651-658 [DOI: 10.1093/plankt/23.6.651].

[4]. Borst, A. and Teunissen, F. E., (1999), "Information theory and neural coding," Nature Neuroscience, Vol. 2, No. 11, pp. 947-957 [DOI: 10.1038/14731].

[5]. Burkitt, A. N., (2006), "A review of the integrate-and-fire neuron model: I. homogeneous synaptic input," Biological Cybernatics, Vol. 95 , No.1, pp. 1-19 [DOI: 10.1007/s00422-006-0068-6].

[6]. Burkitt, A. N., (2006), "A review of the integrate-and-fire neuron model: II. inhomogeneous synaptic input and network properties, "Biological Cybernatics, Vol. 95, No. 2, pp. 97-112 [DOI: 10.1007/s00422-006-0082-8].

[7]. Choudhary, S. K. and Singh, K., "Spiking activity of a LIF neuron in distributed delay framework," International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 3, No. 7, pp. 70-76, June 2016 [DOI: 10.9781/ijimai.2016.3710].

[8]. Choudhary, S. K., Singh, K and Bharti, S. K., "Variability in spiking pattern of leaky integrate-and-fire neuron due to excitatory and inhibitory potentials", 2nd International Conference on "Computing for Sustainable Global Development", pp. 2025-2030, March 2015 [INSPEC Accession Number: 15110012].

[9]. Choudhary, S. K. and Singh, K., "Temporal information processing and stability analysis of the MHSN neuron model in DDF", International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 4, No. 2, pp. 40-45, December 2016 [DOI: 10.9781/ijimai.2016.427].

[10]. Deco, G., Schumann, B., (1997), "Information transmission and temporal code in central spiking neuron," Physical Review Letters, Vol. 79, No. 23, pp. 4697-4700 [DOI: 10.1103/PhysRevLett.79.4697].

[11]. Dilek, S., Cakir, H. and Aydin, M., (2015), "Applications of artificial intelligence techniques to combating cyber crimes: a review", International Journal of Artificial Intelligence & Applications, Vol. 6, No. 1, pp. 21-39 [DOI: 10.5121/ijaia.2015.6012].

[12]. Ermentrout, G. B. and Terman, D. H., (2010), "Mathematical foundations of neuroscience," Springer [DOI: 10.1007/978-0-387-87708-2].

[13]. Fairhall, A., Brown, E. C. and Barreiro, A. (2012), "Information theoretic approaches to understanding circuit function," Neurobiology, Vol. 22, pp. 653-659 [DOI: 10.1016/j.conb.2012.06.005].

[14]. Frank, T. D., (2005), "Nonlinear fokker-planck equations: fundamentals and applications," Springer [DOI: 10.1007/b137680].

[15]. Gerstner, W. and Naud, R., (2009), "How good are neuron models?" Science, Vol. 326, No. 5951, pp. 379-380 [DOI: 10.1126/science.1181936].

[16]. Gerstner, W., Kreiter, A. K., Markram, H. and Herz, A. V., (1997), "Neural code: firing rate and beyond," Proceedings National Academy of Science, USA, Vol. 94, pp. 12740-12741 [PMCID: PMC34168].

[17]. Gerstner, W. and Kistler, W. M., (2002), "Spiking neuron models: single neurons, populations, plasticity," Cambridge University Press [ISBN-10: 0262514206].

[18]. Haykin, S., Principe J. C., Sejnowski, T. J. and McWhirter, J., (2007), "New directions in statistical signal processing: from systems to brains," The MIT Press [ISBN: 9780262083485].

[19]. Higham D. J., (2001), "An algorithmic introduction to numerical simulation of stochastic differential equations," SIAM Review, Vol. 43, No. 3, pp. 525-546 [DOI: 10.1137/S0036144500378302].

[20]. Karmeshu, Gupta, V. and Kadambari, K. V., (2011), "Neuronal model with distributed delay: analysis and simulation study for gamma distribution memory kernel," Biological Cybernatics, Vol. 104, pp. 369-383 [DOI: 10.1007/s00422-011-0441-y].

[21]. Kovacic, G., Tao, L., Rangan, A. V. and Cai, D., (2006), "Fokker-planck description of conductance based integrate-and-fire neuronal networks," Physical Review E, Vol. 20, pp. 021904-1-15 [DOI: 10.1103/PhysRevE.80.021904].

[22]. Lim, H., Kornijcuk, V., Seok, J. Y., Kim S. K., Kim, I, Hwang, C. S. and Jeong D. S. Reliability of neuronal information conveyed by unreliable neuristor based leaky integrate-and-fire neurons: a model study", Nature: Scientific Reports 5, 5, Article number: 9776, 2015. [DOI: 10.1038/srep09776].

[23]. Lo, C. F. and Chung, T. K., (2006), "First passage time problem for the ornstein-uhlenbeck neuronal model," Neural Information Processing: Lecture Notes in Computer Science, Vol. 4232, pp 324-331[DOI: 10.1007/11893028_127].

[24]. Maass, W. "Noise as a resource for computation and learning in networks of spiking neurons", Proceedings of the IEEE, Vol. 102, No. 5, 2014, pp. 860-880 [DOI: 10.1109/JPROC.2014.2310593].

[25]. Rabinovich, M. I., Varona P., Selverston A. I. and Abarbanel H. D. I., (2006), "Dynamical principles in neuroscience," Reviews of Modern Physics, Vol. 78, No. 4, pp. 1213-1265 [DOI: 10.1103/RevModPhys.78.1213].

[26]. Rieke, F., Warland, D., steveninck R. R. and Bialek, W. (1997), "Spikes: exploring the neural codes," The MIT Press [ISBN-10: 0262681080].

[27]. Sejnowiski, T. J., Koch, C. and Churchland P. S., (1998), "Computational neuroscience," Science, Vol. 241, pp. 1299-1306 [PMID: 3045969].

[28]. Sinanovic, S. and Johnson, D. H., (2007), "Toward a theory of information processing," Signal Processing, Vol. 87, pp. 1326-1344 [DOI: 10.1016/j.sigpro.2006.11.005].

[29]. Staub, S., Karaman, E., Kaya, S., Karapinar, H. and Guven, E., (2015), "Artificial neural network and agility", World conference on technology, innovation and entrepreneurship, Procedia – Social and Behavioral Sciences, Vol. 195, pp. 1477-1785 [DOI: 10.1016/j.sbspro.2015.06.448]

[30]. Sterratt, D., Graham, B., Gillies, A. and Willshaw D., (2011), "Principles of computational modeling in neuroscience," Cambridge University Press [ISBN: 9780521877954].

[31]. Tuckwell, H. C., (1988), "Introduction to theoretical neurobiology: volume 1 linear cable theory and dendritic structure", Cambridge University Press [ISBN-10: 0521022223].

[32]. Yang, G. J., Murray, J. D., Repovs G., Cole, M. W., Savic, A., Glasser M. F., Pittenger C., Krystal, J. H., Wang, X. J., Pearlson, G. D., Glahn D. C. and Anticevic A. (2014), "Altered global brain signal in schizophrenia," PNAS, Vol. 111, No. 20, pp. 7438-7443 [DOI: 10.1073/pnas.1405289111].

[33]. Walczak, S. ,(2016) , " Artificial neural networks and other AI applications for business management decision support", International Journal of Sociotechnology and Knowledge Development, Vol. 8, Issue: 4, Article:1, pp. 1-20 [DOI: 10.4018/IJSKD.2016100101].

Author's Biography

Saket Kumar Choudhary: Saket Kumar Choudhary obtained his master degrees in Mathematics from the University of Allahabad, Allahabad, India in 2005, Master of Computer Application (MCA) from UPTU, Lucknow, India in 2010, Master of Technology (M.Tech) from Jawaharlal Nehru University, New Delhi, India in 2014. He is Ph.D (Computer Science and Technology) School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi, India. Currently, he is working as a Senior Data Scientist in Impel Labs. Pvt. Ltd. Bengaluru. His research interest includes mathematical modeling and simulation, dynamical systems, computational neuroscience: modeling of single and coupled neurons, computer vision, digital image processing, machine learning and artificial intelligence.

Dr. Sunil Kumar Bharti: Dr. Sunil Kumar Bharti receives Master's degree (Master of Computer Application) and Ph.D from School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi, India. His primary research area are in Modeling and Simulation, Computational (Neuroscience), image processing and computer vision. He has published various research papers (in reputed journals, international and national conferences.

How to Cite

Choudhary, Saket Kumar and Bharti, Sunil Kumar, "Information Processing in Neuron with Exponential Distributed Delay*",International Journal of Machine Learning and Networked Collaborative Engineering*, Vol. 02 No. 02, 2018,pp 58-66. doi: https://doi.org/10.30991/IJMLNCE.2018v02i02.003.

IJMLNCE JOURNAL

International Journal of Machine Learning and Networked Collaborative Engineering

Journal Homepage: http://www.mlnce.net/home/index.html

DOI : https://doi.org/10.30991/IJMLNCE.2018v02i02.004

Improvement of Automated Learning Methods based on Linear Learning Algorithms

Farzad Kiani

Department of Computer Engineering, Engineering and Natural Sciences Faculty,

Istanbul Sabahattin Zaim University, 34303, Kucukcekmece, Istanbul/Turkey

Farzad.kiani@izu.edu.tr, https://orcid.org/0000-0002-0354-9344

Abstract

In recent years, the learning methods are converted to one of the new research area. These researches are divided into two general categories. The first category recognizes the principles of learning the living entities and its stages. The second is learning based methodology to any machines that the proposed method of this paper is based on it. Learning is defined as changes made in the performance of a system based on experiences. An important feature of learning systems is the ability to improve their efficiency over time. In mathematical terms, it can be stated that the purpose of a learning system is to optimize a task that is not well-known. Therefore, an approach to this problem is to reduce the goals of the learning system to an optimization problem. So, it is defined on a set of parameters and its purpose is to find the optimal set of parameters. In many of the issues raised, there is no knowledge of the correct answers to the problem in supervised learning based methods especially. For this reason, the use of a learning method called reinforcement learning has been considered. The main advantage of this technique over other learning methods is the need for no information from the environment (except amplification signal). The other learning methods as supervised or unsupervised are not appropriate to these problems. In this method, each agent decides the next its actions based on current k-actions instead of one action. In this paper is proposed a new approach based on the reinforcement learning technique that has three versions in order to implementation in different areas. It behaviors based on reward and penalty model. The effectiveness of these interactions with the environment is evaluated by the maximum and minimum of the number of rewards and penalties that are taken from the environment. The three versions are simple, sequential and unstructured linear learning methods so they evaluated in different possibilities to get the appropriate responses. Depending on the needs of any system, they can be used. The mode of convergence of actions in the proposed automaton (machine) in six different scenarios is examined.

Keywords

Automated learning,

linear learning,

smart systems,

reinforcement learning

1 Introduction

In recent years, the process of learning creatures is converted to one of the new research area [1]. These researches are divided into two general categories. The first category introduces learning the live stage presence and recognizes their principles. Aim of the second group is provide a methodology for placing these principles in a machine. Learning is defined as changes made in the performance of a system based on experiences [2]. An important feature of learning systems is the ability to improve their efficiency over time. The most prominent features of learning-based systems are that they improve themselves over time. In mathematical terms, it can be stated that the purpose of a learning system is to optimize a task that is not well known [3, 4]. Therefore, an approach to this problem is to reduce the goals of the learning system to an optimization problem, which is defined on a set of parameters and aims to find a set of optimal (appropriate) parameters. In many of the issues raised, there is no knowledge of the correct answers to the problem, which is required by supervised supervision. For this reason, the use of a learning method called Reinforcement Learning (RL) has been considered. This category of learning method is an orthogonal approach to solving different and more difficult problems. It uses a combination of dynamic programming and supervisory learning to achieve a powerful machine learning system. In the RL is defined a goal for the learning agent so the agent must be achieve it. Consequently, the agent learns how to achieve the target by different tests in various environment [5, 6]. In the RL, a learner's factor in learning through repeated interactions with the environment leads to an optimal control policy. The effectiveness of these interactions with the environment is evaluated by the maximum (minimum) of the number of rewards (penalty) taken from the environment. The main advantage of the RL over other learning methods is the need for no information from the environment (except amplification signal) [2].

The first attempts to use the learning automata in control applications were carried out by Fu et al [7]. In these years, this field has been very intensive and relevant studies. Among these research studies, we can mention the uses of learning automata in the estimation of parameters, pattern recognition, linear update methods, game theory [8, 9], McLaren [10], etc. One of the learning based approaches that can be useful in different environments is Bayesian networks. Bayesian methods provide a standard for optimal decision-making, although in many cases it is incalculable. The [11] proposes A novel decentralized decision making scheme was proposed in [11] that it is based on the Goore Game. The nature of each decision-maker is naturally Bayesian and avoids the difficulty of updating hyperactive parameters of the sibling conjugate primers and calculating them from random sampling from these posterior. Learning-based systems have become very popular and are nowadays used in many business and scientific fields. On the other hand, they have the ability to integrate with todays technologies and research area. On many occasions, the combination of highly valuable products and research is also achieved as multidisciplinary.

A learning automaton consists of two main parts. One of them is stochastic learning automaton that has limited actions number and a random environment in which the automaton is associated with it. The other is the learning algorithm that the automata learns the optimal operation using it. A random automaton is defined as $SA \equiv \{\alpha, \beta, F, G, \phi\}$. Where, α shows set of automata actions. α is defined from index one to r, so, r is number of automatic actions. $\alpha \equiv \{\alpha_1, \alpha_2, ..., \alpha_r\}$. β is input sets of automata that its domain is assumed as m so $\beta \equiv \{\beta_1, \beta_2, ..., \beta_m\}$. F is new status generator function that is defined as $F \equiv \phi \times \beta \rightarrow \phi$. G is output function that the output function that maps the current state to the next output ($G \equiv \phi \rightarrow \alpha$). ϕ presents the input states set of the automata is at the n moment as $\phi(n) \equiv \{\phi_1, \phi_2, ..., \phi_k\}$. The set α includes automated outputs (actions), in which the automation in each step chooses an operation of r for this set to apply to the environment. If the mapping F and G are definite, the automata is called deterministic automata. When the F and G maps are random, the automata is called non-deterministic automata.

Learning automata are divided into two groups of fixed and variable structure automata. In stochastic automata with a fixed structure, the probabilities of automated operations are constant. While in stochastic automata with a variable structure, the probabilities of automated operations are updated in each repetition. In this structure, changing the likelihood of actions is done based on the learning algorithm and the internal state of the automata is represented by the probabilities of the operation of the automata. In fact, each automaton machine has some states as input and output that they are equivalent in almost times. The action probability vector of the operation that is defined at the follow equation defines the internal state of the automaton at the instant n.

$$P(n) = \{p_1(n), p_2(n), ..., p_r(n)\} \tag{1}$$

So that, at the beginning of the activity of the automaton, the probability of its operation is equal and equal (r is the number of automatic operations).

$$\sum_{i=1}^{r} p_i(n) = 1, \ \forall n, \ p_i(n) = Prob[\alpha(n) = \alpha_i] \quad (2)$$

The environment can be represented as $E \equiv \{\alpha, \beta, c\}$. Where, α shows the input sets of environment, β presents the output sets of environment, and c introduces set of penalty probabilities. $\alpha=\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$, $\beta=\{\beta_1, \beta_2, \ldots, \beta_m\}$, $c=\{c_1, c_2, \ldots, c_r\}$.

The input of the environment is one of the r automata actions. β specifies the output (response) of the environment to each action. For example, the system is called P-Model system if the response of the β is binary. In such an environment, $\beta_i(n)=1$ is as an unfavorable or failure response. The favorable or successful answer is when the $\beta_i(n)$ is zero. Other environment model is Q-Model. In this model, $\beta_i(n)$ contains a limited number of values in the interval [0, 1]. In addition, in S-Model, $\beta_i(n)$ is a random variable in the interval [0, 1]. As mentioned above, c specifies the probabilities of penalty (failures) of environmental responses and is defined as equation 3.

$$C_i = Prob\{\beta(n) = 1 \mid \alpha(n) = \alpha_i\}, \ i = \{1, 2, \ldots, r\} \quad (3)$$

It shows that α_i may receive an undesirable response from the environment. The values of α_i are unspecified and is assumed the C_i have at least one unique value. The same way, the environment can be represented by the set of reward probabilities (success). So, it is shown by d_i. The d_i indicates the probability of receiving the desired response to the action of α_i. In static environments, the probability of penalty of α_i are constant. While in non-stationary environments, the probabilities of fines change over time. The connection of random automata with the environment is shown in figure 1.

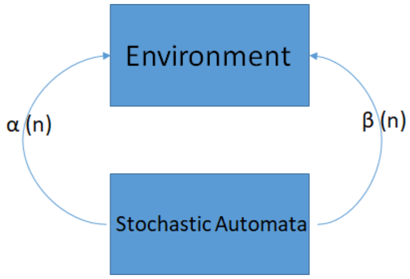


Fig. 1: SLA Schema

The figure 1 refers a learning algorithm that is called the Stochastic Learning Automata (SLA) [10]. Similarly, SLA can be defined as LA={α,β, p, T}. Where, α is defined from index one to r, so, r is number of automatic actions. $\alpha=\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$. β is input sets of automata that its domain is assumed as r so $\beta=\{\beta_1, \beta_2, \ldots, \beta_r\}$. P shows action probability vector and is represented as $p=\{p_1, p_2, \ldots, p_r\}$. In this case, our learning algorithm is:

$$T = p(n+1) = T[\alpha(n), \beta(n), p(n)] \quad (4)$$

If T is a linear operator, the reinforcement learning algorithm is called linear. Otherwise, it is called Non-linear. If the learning automata in repeat n, chooses a sample action such α_i and receive an appropriate response from the environment, the likelihood of α_i action increases and the likelihood of other actions decreases. Conversely, if the response of the environment is undesirable, the likelihood of action of α_i decreases and the probability of other automated actions increases. We can increase the performance of the system by selecting more than one action at the same time while in the previous methods was selected a single action at each stage.

The learning based machines can be applied on the various application area such as smart systems, wireless ad-hoc networks, pervasive systems, IoT, emotional analysis systems etc. [12-14]. In the literature, many learning based techniques are applied on the above applications especially on wireless sensor networks. One the popular used models is cellular automata. It was introduced by Von Neumann at 1940 and after that, he proposed a mathematician named Olam Modelli to study the complex systems behavior. The cellular automata are, in fact, discrete dynamic systems whose behavior is based entirely on local communication. From the point of view of pure mathematics, they can be considered as a branch of topological dynamics and from the point of view of electrical engineering, iterative arrays. In cellular automation, space is defined as a network that is called a cell to each house. Time runs out discretely and its rules are global. Through each cell, each cell acquires its new status by considering its neighbors. Cellular automata can also be considered as computational systems that process the information they encode on their own. In addition, a cellular automaton with its control unit can be interpreted as a SIMD machine. The rules of cellular automation explains how to affect the cell's adherence to its neighboring cells. We call a cell is a neighbor of another cell, if it can affect it in one-step and in accordance with the rules. The rules are generally divided into three categories in this automaton. Firstly is general rules; in this rule, the amount of a cell in the next stage depends on the amount of single neighboring cells in the current state. In the second rule (totalistic Rules), the amount of a cell in the next step depends on the number of neighboring cells that are in different states. In this type of rule, unlike general rule, attention is not paid to single cells. In the last rule (outer totalistic rules), determining the next state of the cell is also effective in the current state.

In the section 2 of the paper is introduced a multi actions and learning based machine that has three versions. So, the evaluation results of them are discussed in the section 3. Finally, the conclusion of the paper is explained in the section 4.

2 Proposing Linear Learning Algorithms

In the standard learning automaton, an action is selected at each step. Then, the probability vector of the actions is updated with considering the appropriate or unfavorable environmental responses. If the appropriate or inappropriate environmental response is not based on the effect of an action, gives unsuccessful results. In this paper is proposed a new type of learning automaton that in this type of machine, instead of choosing one action, k-actions are chosen and then the environment response is received. It is an approach based on reward and penalty system so it is based on reinforcement learning and game theory. The environmental response can be determined by different methods. In the majority voting technique, the optimal appropriate environment response is appropriate when the effect of k/(2 + 1) of the selected action in the environment has the same environment response. Otherwise, the environmental response is inappropriate. All selected actions will be penalized if the response is inappropriate. If the response is desirable, the actions whose effects on the environment have the same environmental response (effective actions in giving the desired environment response) are rewarded and other actions are penalized. The initialized parameters of proposed machine is defined as SLA and is based on Eq. 4.

In this machine, in each step k actions are selected and applied in the environment. The acceptable answer sets of automata to choice of h^{th} action in step of s is calculated based on Equation 5.

$$\alpha^h(s) \sqsubseteq \alpha, r_s^h \leq r \quad \alpha^h(s) = \left\{\alpha 1, \ldots, \alpha_{r_s^h}\right\}, \quad \alpha^1(s) = \alpha \tag{5}$$

In this status, the h^{th} selected answer is deleted from the next actions of the related answer sets. In this machine, three types of ***linear learning algorithms*** are suggested for the automata.

The <u>*first learning algorithm*</u>, which we call the simple linear learning algorithm and is calculated by Equations 6 and 7. So, The Eq. 6 is used for finding appropriate answers and the Eq. 7 is applied to find the inappropriate answers. In here is assumed the h^{th} selected action in n^{th} step is action α_i.

$$p_i^{h+1}(n) = p_i^h(n) + \alpha[1 - p_i^h(n)] \quad and \quad p_j^{h+1}(n) = (1-\alpha)p_j^h(n) \tag{6}$$

$$p_i^{h+1}(n) = (1-b)p_i^h(n) \quad and \quad p_j^{h+1}(n) = \left(\frac{b}{r-1}\right) + (1-b)p_j^h(n), \ \forall j \ \ j \neq i \tag{7}$$

The relationship between these two equations with the probability vector is expressed by equation 8.

$$P_i^1(n) = p_i(n) \quad \forall i \quad such\ that\ p_i \in p(n)$$

$$p_i(n+1) = p_i^{k+1}(n) \quad \forall i \quad such\ that\ p_i \in p(n+1) \tag{8}$$

As with the relationships of the first learning algorithm, in this type of learning algorithm, at each stage, all probability vector members are updated to a number of k times.

The *second learning algorithm* is called a sequential linear learning algorithm and is formulized by Equations 9 and 10. So, The Eq. 9 is used for finding appropriate answers and the Eq. 10 is applied to find the inappropriate answers. In here is assumed the h^th^ selected action in n^th^ step is action α_i.

$$\hat{p}_i^{h+1}(n) = \hat{p}_i^h(n) + a\left[1 - \hat{p}_i^h(n)\right] \quad and \quad \hat{p}_j^{h+1}(n) = (1-a)\hat{p}_j^h(n) \quad \forall j \;\; j \neq i \qquad (9)$$

$$\hat{p}_i^{h+1}(n) = (1-b)\hat{p}_i^h(n) \quad and \quad \hat{p}_j^{h+1}(n) = \left(\frac{b}{r-1}\right) + (1-b)\hat{p}_j^h(n) \quad \forall j \;\; j \neq i \qquad (10)$$

The value of $\hat{p}$ is calculated by Equation 11.

$$\hat{p}_i^h(s) = \frac{p_i^h(s)}{M} \quad \forall i \;\; such\ that \;\; \alpha_i \in \alpha^h(s), \quad M = \sum_{j=1}^{r_s^h} p_j^h(s) \quad \forall j \;\; such\ that \;\; \alpha_j \in \alpha^h(s) \qquad (11)$$

In the Eq. 9 and 10, α presents reward parameter and b shows the penalty parameter. The probability vector of the automata is updated according to the Equation 12 in order to return the machine to normal conditions and select the (h+1)^th^ action.

$$p_i^{h+1}(s) = \hat{p}_i^{h+1}(s).M \quad \forall i \;\; such\ that \;\; \alpha_i \in \alpha^h(s)$$

$$p_i^{h+1}(s) = p_i^h(s) \quad \forall i \;\; such\ that \;\; \alpha_i \notin \alpha^h(s) \qquad (12)$$

As it is known from the relations of the second learning algorithm, at each step of in this type, the probability of the action of the h^th^ selected action will be updated for h times, and actions that are not selected are updated to k times.

The *third learning algorithm* is called unstructured linear learning algorithm and is formulized by Equations 13 and 14. So, The Eq. 13 is used for finding appropriate answers and the Eq. 14 is applied to find the inappropriate answers. In here is assumed the h^th^ selected action in n^th^ step is action α_i.

$$\hat{p}_i^{h+1}(n) = \hat{p}_i^h(n) + a\left[1 - \hat{p}_i^h(n)\right] \quad and \quad \hat{p}_j^{h+1}(n) = (1-a)\hat{p}_j^h(n) \quad \forall j \;\; j \neq i \qquad (13)$$

$$\hat{p}_i^{h+1}(n) = (1-b)\hat{p}_i^h(n) \quad and \quad \hat{p}_j^{h+1}(n) = \left(\frac{b}{r_s^h-1}\right) + (1-b)\hat{p}_j^h(n) \quad \forall j \;\; j \neq i \qquad (14)$$

The value of $\hat{p}$ is calculated by Equation 15. Where, $\alpha_k(s)$ presents k actions selected in s step. In addition, α presents reward parameter and b shows the penalty parameter.

$$\hat{p}_i^h(s) = \frac{p_i^h(s)}{M} \quad \forall i \;\; such\ that \;\; \alpha_i \in \{\alpha - \alpha_k(s) + \alpha_i\}, \quad M = \sum_{j=1}^{r_s^h} p_j^h(s) \quad \forall j \;\; such\ that \;\; \alpha_j \in \{\alpha - \alpha_k(s) + \alpha_i\}$$

(15)

The probability vector of the automata is updated according to the Equation 16 in order to return the machine to normal conditions and select the (h+1)^th^ action.

$$p_i^{h+1}(s) = \hat{p}_i^{h+1}(s).M \quad \forall i \;\; such\ that \;\; \alpha_i \in \{\alpha - \alpha_k(s) + \alpha_i\}$$

$$p_i^{h+1}(s) = p_i^h(s) \quad \forall i \;\; such\ that \;\; \alpha_i \notin \{\alpha - \alpha_k(s) + \alpha_i\} \qquad (16)$$

As it is known from the relations of the third learning algorithm, at each step of in this type, the probability of the selected actions will be updated for one time, and actions that are not selected are updated to k times.

3 Evaluation

With a sample of automata, we examine the mode of convergence of operations in this type of automaton (machine) in MATLAB software simulation. This sample machine has 6 actions, and each step is selected 3 steps. By responding to the environmental responses (responses from the effects of each action on the environment), the desired response is detected. Given the appropriate response, the related actions are rewarded and penalized. The probability of receiving the appropriate response from the environment for each operation is shown in Table 1.

Table 1. Possibility to receive the optimal appropriate responses from the environment for each of the actions in the proposed learning machine.

Action #	1	2	3	4	5	6
Possibility to get the appropriate response	0.1	0.9	0.1	0.7	0.7	0.5

The reward and penalty parameters are set at 0.01. The mode of convergence of automated actions for different learning algorithms is seen in Figures 2,3 and 4.

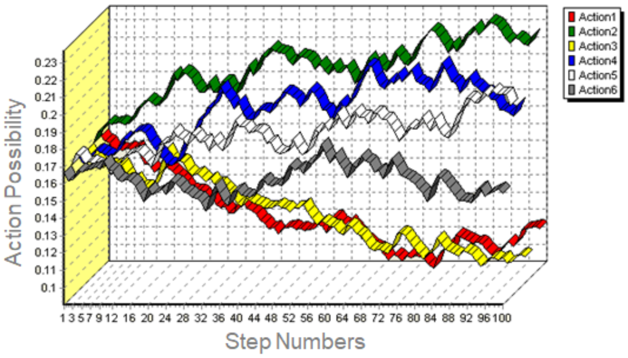


Fig. 2: Results of convergence of automated actions in simple linear learning algorithm.

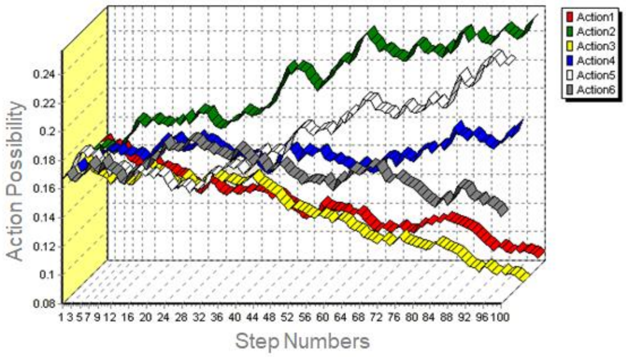


Fig. 3: Results of convergence of automated actions in sequential linear learning algorithm.

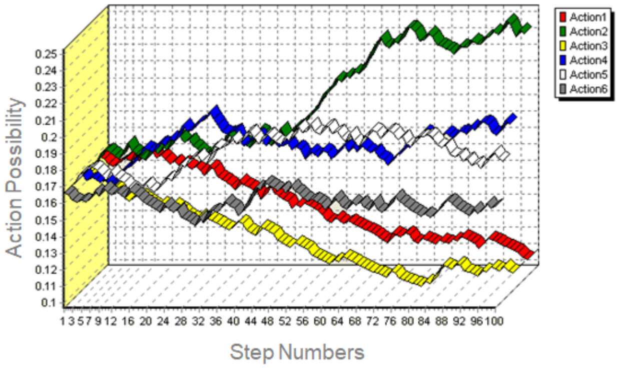


Fig. 4: Results of convergence of automated actions in unstructured linear learning algorithm.

It is possible to use this method in various fields such as wireless ad-hoc networks and IoT based on the results. Each proposed algorithm can be useful in the related environment application based on system requirements. In additional, this method can be used as a mixed approach in different learning based algorithms as Bayesian networks, game theory techniques, Q-learning, Sarsa, neural network and deep learning.

4 Conclusion and Future Works

In this paper was proposed a new model of learning automaton machine that it worked based on K-actions contrary to existing methods in the literature. It affects the response of the environment based on selecting and analyzing the k-actions. In this paper was introduced three different linear learning methods as simple, sequential and unstructured. The Analysis results on the six samples of actions, which the k was supposed 1 to 6, are shown the each method's performance. They can be used in various application areas that we will apply it on the wireless Ad-hoc networks and IoT research area. The third method seems to be high performance on wireless networks that it is one of the future works. It is possible to use this method in various fields such as wireless ad-hoc networks and IoT based on the results. Each proposed algorithm can be useful in the related environment application based on system requirements. In additional, this method can be used as a mixed approach in different learning based algorithms as Bayesian networks, game theory techniques, Q-learning, Sarsa, neural network and deep learning. In future work, we have a plan in research study on wireless sensor networks based on the proposed algorithms. We will analyze the simulation results and will compare them with the relevant methods in the literature.

References

[1]. V. Saritha, P. V. Krishna, S. Misra and M.S. Obaidat, "Learning Automata based Optimized Multipath routing using Leapfrog Algorithm for VANETs", IEEE ICC 2017 Mobile and Wireless Networking, 1-5, 2017. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7997401

[2]. K. S. Narendra and M. A. L. Thathachar, "Learning automata: An introduction", *Proceedings of the Prentice Hall*, 1989. https://dl.acm.org/citation.cfm?id=64802

[3]. Z. Shariyat, A. Movaghar, M. Hoseinzadeh, "A learning automata and clustering-based routing protocol for named data networking", *Telecommunications Systems*, 65(1), 9-29, 2017. https://link.springer.com/article/10.1007/s11235-016-0209-8

[4]. H. Ge and Sh. Li, "A Parameter-Free Learning Automaton Scheme", *Cornell University Library*, 1-13, 2017. https://arxiv.org/pdf/1711.10111.pdf

[5]. E. Mance and S. H. Stephanie, "Reinforcement learning: A tutorial", *Proceedings of the Wright Laboratory*, 1996. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.33.2480&rep=rep1&type=pdf

[6]. K. Arulkuraman, M. Peter, M. Brundage and A. Bharath, "Deep Reinforcement Learning: A brief survey", *IEEE Signal Processing Magazine*, 26-38, 2017. https://ieeexplore.ieee.org/document/8103164/

[7]. K. S. Fu and T. J. Li, "Formulation of learning automata and automata games", Proceedings of the Information Science, 1(3), 237-256, 1969. https://www.sciencedirect.com/science/article/pii/S0020025569800101

[8]. C. S. Chasparis, "Stochastic Stability of Perturbed Learning Automata in Positive-Utility Games", 1-16, 2018, https://arxiv.org/pdf/1709.05859.pdf

[9]. A. Jitpattanakul, "Learning k-edge Deterministic Finite Automata in the Framework of Active Learning", *International Journal of Applied Engineering Research*, *International Journal of Applied Engineering Research*, 12(6), 6050-6054, 2017. https://pdfs.semanticscholar.org/ef37/f0bf558148dd2b830abdd4dfbb406b50e2f7.pdf

[10]. R. W. McLaren, "A stochastic automaton model for synthesis of learning systems", Proceedings of the IEEE Transactions on System Science and Cybernetics, 2, 109-114, 1966. https://link.springer.com/chapter/10.1007/978-1-4615-9050-7_5

[11]. O. Christoffer, S. Glimsdal, "Accelerated Bayesian learning for decentralized two-armed bandit based decision making with applications to the Goore Game", *Appl Intell, Springer Science and Business Media, LLC 2012,* 1-10, 2012. https://brage.bibsys.no/xmlui/bitstream/handle/11250/137969/Granmo_2012_Accelerated.pdf?sequence=1

[12]. S. Tanwer et al., "LA-MHR: Learning Automata Based Multilevel Heterogeneous Routing for Opportunistic Shared Spectrum Access to Enhance Lifetime of WSN", *IEEE Systems Journal*, 1-11, 2018. https://ieeexplore.ieee.org/document/8351993/

[13]. D. Mendez, I. Papapanagiotou, B. Yang, "Internet of Things: Survey on Security and Privacy", 1-15, 2017, https://arxiv.org/pdf/1707.01879.pdf

[14]. F. Kiani, "Reinforcement Learning Based Routing Protocol for Wireless Body Sensor Networks", International Symposium on Cloud and Service Computing (SC2), 2017 IEEE 7th, 71-78, 2017. https://ieeexplore.ieee.org/abstract/document/8315358/

Author's Biography

Farzad Kiani: Farzad Kiani has PhD degree and he is Assistant Prof. in computer engineering dept. at Istanbul Sabahattin Zaim University (IZU) from 2014 to now. In addition, he is deputy head of computer engineering and head of wireless sensor networks and IoT Labs at IZU from 2016 and 2017 to now, respectively. His current research interests include wireless sensor networks, Vehicle Ad-hoc networks, flying Ad-hoc networks, machine learning, game theory and IoT. He is also working on optimization algorithms and security.

How to Cite

Kiani, Farzad, "Improvement of Automated Learning Methods based on Linear Learning Algorithms*", International Journal of Machine Learning and Networked Collaborative Engineering,* Vol. 02 No. 02, 2018, pp. 67-74. doi: https://doi.org/10.30991/IJMLNCE.2018v02i02.004.

Vol. 02, No. 2 ,(2018), 075–085, **ISSN:2581-3242**

IJMLNCE
JOURNAL

International Journal of Machine Learning and Networked Collaborative Engineering

Journal Homepage: http://www.mlnce.net/home/index.html

DOI : https://doi.org/10.30991/IJMLNCE.2018v02i02.005

IoT with Big Data Framework using Machine Learning Approach

Jyotir Moy Chatterjee

Department of Computer Science &Engineering

GD-RCET,Bhilai,India

jyotirm4@gmail.com, https://orcid.org/ 0000-0003-2527-916X

Abstract

In future IoT (Internet of Things), big data (BD) & machine learning (ML) disclosure for expansive scale modern robotization application significance of mechanical internet is expanding step by step. The interconnection by means of the Internet of computing gadgets installed in simple items, empowering them to send and receive information. BD is informational collections that are so voluminous and complex that traditional data processing application are insufficient to manage them. ML is a subset of artificial intelligence that regularly utilizes measurable procedures to enable PCs to "learn" with information, without being expressly modified. A few differentiated advancements, for example, IoT, computational intelligence, ML, BD & sensor technology can be fused together to enhance the data management & information revelation effectiveness of expansive scale robotization of applications. An expanding measure of significant data sources, propels in IoT & BD advancement, also the accessibility of an extensive variety of ML computations offers new potential to convey logical administrations to industries. In any case, there is as yet a hole in joining the present best in class in an incorporated system that would help lessening improvement costs & empower new types of administrations. Voluminous measures of data have been created, since the previous decade as the scaling down of IoT gadgets increments. However, such data are not valuable without scientific power. Various BD, IoT investigation arrangements have empowered individuals to acquire profitable knowledge into extensive information created by IoT gadgets. However, these arrangements are still in their earliest stages & the domain does not have a thorough review on this. Here we endeavored to give a reasonable more profound understanding about the IoT in BD structure alongside its different issues & challenges & concentrated on giving conceivable solutions by ML strategy.

Keywords

Internet of Things

Big Data ,

Computational Intelligence,

Sensor Technology,

Machine Learning

1. Introduction

The IoT is that the system of gadgets, automobiles, house apparatuses & various things inserted with hardware, software, sensors, actuators, & system availableness that empowers these articles to associate & trade knowledge. Each factor is awfully recognizable through its inserted problem-solving framework nonetheless will between work within the present web foundation. Specialists value that the IoT can consist of about thirty billion queries by 2020. It's likewise evaluated that the worldwide market estimation of IoT can reach $7.1 trillion by 2020. The IoT permits articles to be detected or controlled remotely crosswise over existing system framework, creating open doors for additional simple mixture of the physical world into computer primarily based frameworks, & conveyance regarding increased proficiency, exactitude & money advantage nevertheless lessened human intercession. At the purpose once IoT is enlarged with sensors & actuators, the technology turns into an occurrence of the additional broad category of digital physical frameworks, that in addition envelops advances, for instance, sensible grids, virtual power plants, sensible homes, intelligent transportation & sensible cities [1]. "Things", within the IoT sense, will touch to a good type of gadgets, for instance, heart checking implants, silicon chip transponders on cultivate creatures, cameras gushing live nourishes of untamed creatures in city district waters, cars with worked in sensors, deoxyribonucleic acid examination gadgets for ecological/sustenance/pathogen perceptive, or field operation gadgets that facilitate firefighters in hunt & save operations. Lawful researchers suggest with relation to "things" as associate "inseparable mix of apparatus, software, knowledge & administration". The expression "the web of things" was began by Kevin Sir Frederick Ashton of Procter & Gamble, later MIT's Auto-ID Center, in 1999. For instance- The Philips Hue Connected bulb enables the client to wirelessly control the lighting framework in the home, & comprises of an Ethernet empowered bridge that acknowledges orders from the client application & conveys these to the bulbs utilizing the ZigBee-Light link protocol. The information trade between the application & the bridge is by means of HTTP summons & isn't encoded, so a busybody can without much of a stretch find the activities the client performs on the bulb. Further, despite the fact that the gadget actualizes get to control as a white-listed arrangement of clients, this rundown can be separated by any attacker, who would then be able to take on the appearance of an authentic client, consequently picking up control over the bulb [2]. From a development perspective, the most essential thing is to get occasions from IoT-related devices. The contraptions can be related with the framework using Wi-Fi, Bluetooth, or other gadgets having the ability to send messages to a specialist using some especially portrayed tradition.

A champion among the most conspicuous & by large used traditions is Message Queue Telemetry Transport (MQTT). Mosquitto is an outstanding open-source MQTT authority. IoT & BD on a very basic level are two sides of a comparable coin. Directing & expelling a motivating force from IoT data is the best test that associations stand up to. Associations should set up a proper examination arrange/system to explore the IoT data. Moreover, they should recall that not all IoT data is basic [3]. A fitting examination arrange should be established on three parameters: execution, right-measure system, & future advancement. For execution, an uncovered metal server, a lone inhabitant physical server focused on a single customer, is the best fit. For system & future improvement, half & half is the best approach. Half & half plans, which include cloud, directed encouraging, colocation, & gave encouraging, join the best highlights from various stages into a lone perfect condition. Multiple Service Providers (MSPs) are also managing their phases to manage IoT data. MSP vendors are routinely tackling the establishment, execution, & devices side to cover the entire IoT space. An IoT device creates relentless surges of data adaptable & associations must manage the high volume of stream data & perform exercises on that data. The exercises can be event relationship, metric computation, bits of knowledge availability, & examination. In a conventional tremendous data circumstance, the data isn't by & large stream data, & the exercises are exceptional. Building an examination respond in due order regarding manage the measure of IoT data should be finished in light of these qualifications. Once the data is gotten, the obligation of the development organize is to store the IoT data. Various associations use Hadoop & Hive to store gigantic data. Regardless, for IoT data, NoSQL chronicle databases like Apache CouchDB are more suitable in light of the fact that they offer high throughput & low idleness. These sorts of databases are development less, which supports the flexibility to incorporate new event creates easily. Other surely understood IoT contraptions are Apache Kafka for widely appealing message encouraging & Apache Storm for consistent stream taking care of [12]. There is no reasonable definition for BD [3]. It is defined in view of a portion of its qualities. The BD does not mean the size. There are three attributes that can be utilized to define BD, as otherwise called 3V's: volume, variety & velocity[4].

Volume identifies with size of the information, for example, terabytes (TB), petabytes (PB), zettabytes (ZB), & so on. Assortment implies the sorts of information. Likewise, contrast sources will create BD, for example, sensors, devices, social networks, the web, cell phones, & so forth. Speed implies how as often as possible the information is created (e.g. each week, month, year, etc.) [4]. BD will be information sets that are voluminous & muddled. These information sets allow knowledge handling application to manage them. Huge knowledge challenges incorporate catching knowledge, knowledge storage, knowledge investigation, look, sharing, exchange, illustration, questioning, refreshing & knowledge protection. There are 3 unit measurements of huge knowledge - volume, variety & velocity. Recently, the expression "BD" contains a tendency to touch to the employment of discerning examination, shopper conduct investigation, or sure different propelled knowledge investigation methods that concentrate associate incentive from knowledge, & typically to a particular size of information set. "There is no uncertainty that the amounts of information currently accessible area unit evidently vast, but that's not the foremost vital traditional for this new knowledge biological community." Analysis of information sets will discover new relationships to "spot business patterns, avert infections, battle wrongdoing & then on." Scientists, business directors, specialists of drug, commercial enterprise & governments alike habitually meet challenges with immense data-sets in territories as well as web look, fintech, urban information science, & business information science. Researchers provide limitations in e-Science work, & additionally meteorology, genomics, connectomics, interwind material science replications, science & common survey. Subsequently, alongside the advantages we are acknowledging from IoT, there are natural difficulties, both as far as specialized framework & the administration of data & individuals. In spite of late advances in IoT framework catching, putting away, & preparing tremendous measure of heterogeneous sensor data is as trying as ever. The exponential increment in the number & variety of sensors likewise makes it troublesome for IoT experts to adequately plane how to use the data, what sort of administrations to furnish with the data & for whom. BD is portrayed by different highlights & they are in particular volume, variety, velocity & veracity. BD is constantly computed in enormous sums (volume), is an accumulation of heterogeneous data (variety), finds in typically ongoing rate (velocity speed by which data is gotten (veracity). It can be more clarified as [3]-

- Volume- Amount of data produced & put away & the size of data establishes that the data can be called possibly big or not.
- Variety - Implies the data type & its temperament, kind of data etc. This will viably help in acknowledging in what manner can the data be isolated.
- Velocity - It alludes to the speed at which the data is delivered. Data is made & prepared in like manner relying on prerequisite of the data in view of the need & necessity.
- Variability – Non-regularity of the data insights can hamper procedures handling & overseeing it.
- Veracity - The nature of acquired data must be exact & legitimate it can't be a fraudulent. Overseeing such tremendous measure of data turns out to be very troublesome & such a big & testing undertaking.

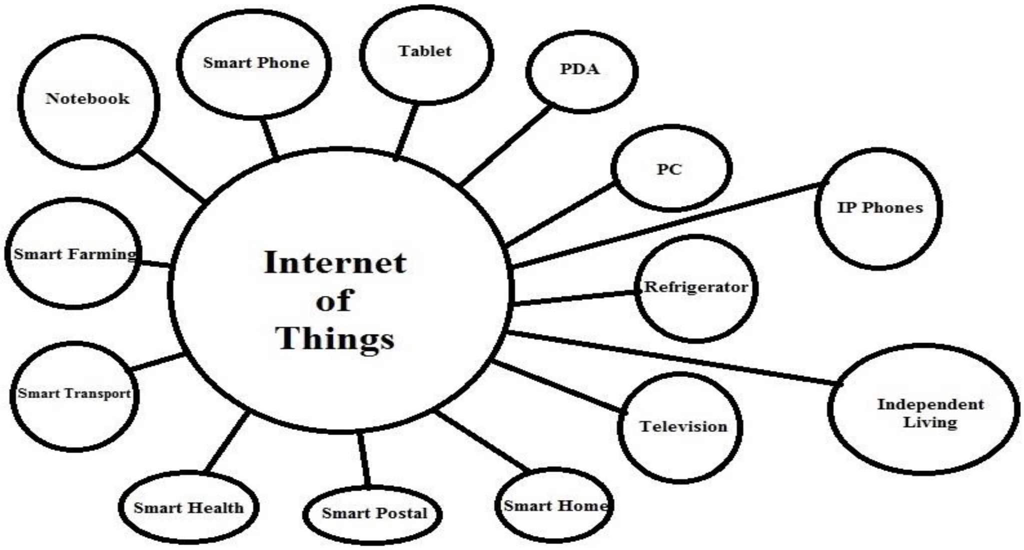


Fig. 1. Internet of Things (IoT)

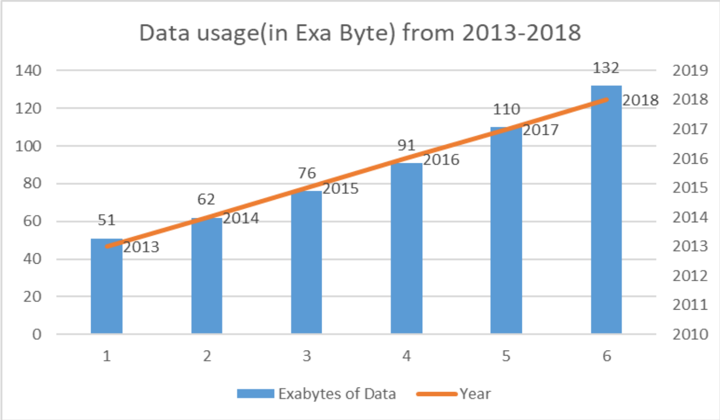


Fig. 2. Per month data usage from 2013 to 2018

2. Applications of IoT

IoT is creating an environment for merging the general populace JOT will be executed wherever where it brings inclinations to individuals. Mobiles now step by step are so valuable & are linking individuals to gadgets more with the advance of innovation. This progress of "IoT "will bring huge open portals, it will in like way reach to point to interface the existing structures & after that review that by connecting more things conceivable as a result of the WSNs & other innovations. Information will then pass on or move from a one area to the other & with this cloud will come into picture & will therefore transport the Data the IoT is propose various range electrical things, programming, sensors, & distinctive framework network to gather & for trading information. With the help of some framework network. With the help of IoT, it will wind up plainly conceivable to perceive & control remote contradictions within live framework configuration, creating chances for things & PC based structures, & proving in update capacity, exactness & cash related good position. Some reliable applications of IoT can occur in Smart urban networks like smart parking i.e. See of the parking space receptiveness in the city. Principal Health & various utilizations in detection of smart telephone with the help of wi-fi or Bluetooth gadget. Smart Lightening & Some extraordinary utilizations in Smart Environment are Forest fire prediction, Ice level monitoring, Pre-Earthquake detection, Water monitoring, Pollution level in the sea, Floods. Other areas where IoT can help making things basic & portable are edge get the opportunity to control, Other domains are air pollution, Smart Grid, Radiations level, Intelligent shopping applications, Flight tracking, Energy & Water use & these different courses there are different upcoming zones where IoT has wound up being blessings [5]. [12] presented a model for IoT based savvy parking framework that endeavors to computerize vehicle parking framework that jumps to limit vehicle parking issue. [13] discussed the idea of 'connected cars' can be utilized to perform 'predictive car maintenance'. [14] unmistakably showed that the handshaking model of IoT and IoC ensures a keen life in near future.

3. Challenges that IoT & Big Data have bought into picture:

Following are some of the challenges that IoT & BD have bought into the picture

- Architecture- The essential concerning architecture is Data centers ready to control this extra stream of heterogeneous data. At the present time, utilized IoT models is that they are intended for the most part little scale IoT structure. This stockpiling of enormous measure of data would require some portion of capability constrain along these lines bringing about High cost.
- Power- This is the most imperative thing for supporting a big & colossal data JOT applications require to keep running for a considerable length of time to lessen the general vitality consumption. Consequently, Power is must. Harvesting power & stretching battery life is required. IoT creeps into the vitality sector.

- Security - This is major at each layer. It is the principal & first need in IoT. Expanded computerization & digitization makes new security concerns. Expectation of information is an issue since two organizations are related with each other. Since with the commercialization of web, security has been taken as a significant stress since this extend to coat different protection issues like individual, cash related exchanges, & the odds of cybertheft. IoT security is constant. various kinds of gadgets that cooperate IoT & the information writes find will refinement in exchange perspective – protocols for correspondence, rough gadgets & this will pass on with itself the danger of information security. This moved measure of information in IoT is new to specialists & they emphatically require learning & experience which will without a doubt increment the danger of security. Any resentful threat can reach to the undermine the information – it can likewise hamper within related gadgets. There must be some worked in control development. Security isn't simply stomach muscle out the product level but instead it's additionally about either the hardware level, programming or the framework layer. There must be security answers for dodge, perceive & react to pernicious direct. At the hardware level, there must be some symmetric cryptography at programming level approval & against cloning is required though at mastermind level IP affirmation is required.
- Complexity - It need to be easy as opposed to being unpredictable. At first it would not be that easy to understand & use certain applications which can help in using these IoT features. There has to be easy setup for customers to understand & use.
- Sensing - Sensing of the gadgets is important. So, there must be expansive sensing innovations. IoT is about sensor's introduced advancement. So, there must be novel approaches to detect & pass on information from this physical universe of information & after that store it to the cloud.
- Creating Knowledge & BD - There will be persistently immense measure of heterogeneous data. It will be amazingly required to thrive methodologies that will adjust or change over this data into usable data the measure of data assembled will be gigantic. Late details tell that Every minute, 300 million messages are sent,3.8 million Facebook likes are created, send's thousand's tweets, & numerous pictures are exchanged to Facebook. By 2020, it's evaluated this number will build billion's so to manage this BD center's inevitable included & so expectedly persistent sensor's future required.
- Data is encapsulated, however not used completely – The majority of loT ventures do incorporate information gathering, yet not a lot of are totally using the open entryway that information gives. Just 17% of survey reports demonstrated that they don't catch data as a component of their loT projects. While larger part, 89%, are working together data, just a small number (6%) report that they are drawing the most of the data by totally catching & examining data in an auspicious fashion. The greater part of participants (58%) are attempting & are doing some analytics, however they know they can improve the situation.
- Cost - There will be a great deal of cost engaged with the same. Cost of the hub actualized, vitality consumed, improvement included & deployment of the same would be a troublesome task.
- Adaptability - Adaptability to condition, adjusting new faults & blunder at dynamic time seems to be an issue. Which has to be dealt with all the more precisely because this can mislead any data.
- Self - Learning – This is the key, because design discovery, auto setup has to be take care.
- Deployment - Deployment of the gadgets to particular regions would incorporate piece of cost additionally it may likewise incite blunders so envisioning such mistakes in the meantime additionally is by all accounts tuff additionally confinement would be incorporated bringing about cost.
- Maintenance - After the deployment of IoT devices keeping up them would be sufficiently troublesome. This would require part of troubleshooting steps as well as repetitive cost

would be included. So, it's just creating the gadget as well as appropriate support is the worry so as to get productive services.

- Privacy - The connection between IoT will give numerous helpful & valuable services for each individual not just to make opportunities but rather will damage certain rules & privacy statements. In future to solve this privacy concerns & policies for the system this will be the significant territory of concern. IOT worldview will have pass on user's request for the data validation & protocols such that its request can be assess against the policies so that they can give or deny access. There is a prerequisite for new protocols & definition because the accompanying requirements can't be expressed from this present scenario [3].

4. IoT's Data Problem

Following are the major IoT's Data Problem

- Nobody will wear 50 devices

On the off chance that there's one lesson the present IoT new companies have gained from their fizzled science venture ancestors, it's that things should be straightforward & turnkey. Therefore, gadgets are intended to complete one thing extremely well. A culmination of this is there's a great deal excessively specialization happening — a gadget particularly, barely intended to gauge rest, or eating rate, or knee wellbeing.

Tragically, nobody can charge, oversee, & wear fifty gadgets, resembling AN unbalanced garage deal machine. VentureBeat's Harrison Weber puzzled out the way to try on fifty-six distinct wearables at CES. With this various contestant, the trade can crash. Wearables these days are a sophisticated knit, AN uncommon interlacing of purpose arrangements endeavoring to hide a personality's life. To accomplish ease, organizations have over-concentrated on a solitary drawback, or a solitary utilize case, dishonorable themselves that their foothold is admittedly a possible market. The aisles of CES were coated with processed yoga mats, shrewd sun sensors, brain disorder locators, & instrumented snowboard ties.

- More inference, less sensing

Consider the previously mentioned sun sensor. Do you truly require a wristband that detects what amount of daylight you've been presented to? Or then again can your cell phone rather measure light levels intermittently (which it does to decide screen splendor at any rate), choose whether you're outside, & check the UV file? The last is inference, instead of sensing, & it's presumably adequate. At the point when the IoT sprawl at last triggers a mass elimination, just a couple of organizations will survive. Huge numbers of the survivors will be the ones that can discover more data by inference, & that implies groups that have an information science foundation.

Early forms of Jawbone's wearable, for instance, requested that wearers log their movement physically. More late forms are more astute: the gadget sees a time of movement, surmises what that action was by contrasting it with known examples — would you say you were playing ball for a half hour? — & utilizes your reaction to either strengthen its figure, or to refresh its aggregate comprehension of what ball feels like.

- Data Mining

This spill of gadgets moreover suggests a spill of learning. Unless you are one in all the vast wearable players — Jawbone, Fitbit, Withings & a humble pack of others — you potentially don't have enough purchaser data to frame imperative achievement revelations concerning your customers' lives. this offers the substantial players a strong first-mover advantage. At the reason once the wearables area emphatically joins together, the greater part of the information that failed associations accumulated will be lost. there is small sharing of data transversely finished item offerings, & admission is frequently a significant comma-disconnected record. consider that one in all the chief grounded reasons individuals don't alteration from Apple to golem is that the shared characteristic of the shopper experience & consequently the substance in iTunes. Correspondingly, inside the IoT world, interfaces & information debilitate trading. Tragically, this implies unfaltering wars over information organizes in AN extraordinary kind of cutting edge gerrymandering — choice it Data Mining — as every businessman racers for position, trying to be the central focus reason for our Eudaimonia, kid raising, home, or finances.

As Samsung, corporate official metal Yoon same in his CES keynote, "I've distinguished individuals say they need to shape a lone working structure for the web of Things, regardless these individuals just work with their own particular devices." Walking CES, you see numerous makers from Shenzhen propelling the building things of the IoT. Headways like surface sensors — that just months back were naturally released from puzzle personnel labs & commended on school sites — would right now be prepared to be had at scale from China. Deterrents to area crumble quick. What remains for IoT associations are thought, apportionment, & learning. At the reason once specific advances break up hack cleave, associations have almost no inspiration to team up on the information they assemble. there isn't any data lake in wearables, just load enviously watched streams.

- Context is everything

In the event that info does not modification your conduct, why strive gathering it? perhaps the best info drawback the IoT faces is corresponding the data it gathers with moves you'll build. consider V1bes, that calls itself a "psyche application." It quantifies feelings of tension & neural structure action. Sociometric Solutions will likewise by taking note of the tone of my voice, & might anticipate my feelings of tension exactly. That sounds valuable: it would be extraordinary to understand however targeted on i used to be at a selected time, or once my mind was typically dynamic. In any case, unless I will see the individual to whom I used to be talking, or hear the words i used to be thoughtful, around then, it's tough to form a move. the data discloses to Pine Tree State i am targeted on; it does not make known to Pine Tree State United Nations agency's setting off my eternal misery or who influences my eyes to illuminate. There is also trust here. within the event that I had a photograph stream of systematically, & with it a voice recorder, I'll have the capability to examine my identity with (& whom to remain away from). New businesses like Narrative Clip, that invariably logs my life by taking a photograph at regular intervals & utilizing calculations to settle on that of these images are intriguing, would possibly offer Pine Tree State some insight concerning what depart my pressure. Also, versatile recorders like Kapture will record discussions with time stamps; their transcripts, dissected, might alter Pine Tree State to examine however I answer specific points [6].

4.1. Data Security Issues

The IoT has given new security challenges that can't be controlled by customary security systems. Confronting IoT security issues require a move. For example, how might you manage a circumstance when the TV and security camera at your home are fitted with cloud Wi-Fi get to. Scarcely any security issues are

- Secure computations in disseminated environment
- Secure data focuses
- Secure transactions
- Secure sifting of redundant data
- Scalable & secure data mining & analytics
- Access control
- Imposing real time security, & so forth.,

A multi-layered security structure and legitimate framework system will help stay away from assaults and shield them from dissipating to various parts of the framework. An IoT structure should take after exhaustive framework get to control systems and then permitted to interface. Software-defined networking (SDN) propels should be used for point-to-point and point-to-multipoint encryption in blend with orchestrate identity and access approaches.

4.2. IOT & ML Approach

ML calculations can be utilized to make forecasts in view of information designs. For instance, in a Mayo Clinic study, action information was associated with recuperation rates for heart patients. [8] some similar ML & prescient calculations are the reason for various associated shrewd shopper gadgets. Nest-thermostats are a case of a gadget that use information examples to foresee the favored temperature in a particular room at a specific time of day. (Another control & streamlining illustration is seen at a collected neighborhood level, where control utilities can move vitality loads at top circumstances by remotely altering—with the inhabitant's consent—hundreds or thousands of Nest gadgets by two or three degrees.) Other customer gadgets incorporate those that gain from voice designs, (for example, Echo, an individual collaborator write gadget from Amazon [9]) to those that gain from considerably more mind boggling conduct & action designs, (for example, Jaguar's Land Rover monitoring system, which "depends on a muddled programming which empowers the auto to study, anticipate, check, & remind the auto's tenants [to] help the driver auto-designate his errands & influence him to focus more on his driving."7). Advancement calculations utilize ML components to use information from the two sensors & wise gadgets that collaborate under unique conditions. These variable conditions can't be unequivocally anticipated past specific parameters. The calculations should detect, react, & adjust. For instance, as autos go up against more obligations from the driver, they will collaborate with more ecological wellsprings of information (sensors, lights, different autos, & so on). Classes of uses in mechanical robotization, coordination & transportation, control network & vitality systems, activity administration, security systems, & other "systems of systems" will give machines a chance to discuss straightforwardly with different machines. Moreover, such applications will enable machines to decipher information streams in view of calculations that can develop & adjust, so the machines can accomplish the coveted end states given certain operational parameters.

5. Big Data Challenge can be solved using ML

ML will become an adult this year, moving from the exploration labs & evidence of-idea usage to forefront business arrangements. Enroute, it will help control developments, for example, autonomous vehicles, exactness cultivating, remedial medication revelation, & propelled extortion location for budgetary foundations. ML meets with insights, software engineering, & counterfeit consciousness, focusing on the improvement of quick & effective calculations to empower continuous data preparing. As opposed to simply take after expressly modified guidelines, these ML calculations gain as a matter of fact, making them a key part of computerized reasoning stages.

5.1. ML helps to handle IoT data flows

ML could likewise facilitate us with a challenge from one amongst a year ago the most hummed concerning innovation advancements: The IoTs. the first of huge information investigation grew up around the stream of information created by web-based social networking, net primarily based looking, online recordings, net surfing, & alternative shopper created online behaviours, as indicated by Vin Sharma, the director of ML arrangements in Intel's information Canter cluster. Breaking down these monstrous datasets needed new advancements, all-mains distributed computing, & virtualization programming, as an example, Apache Hadoop & Spark. It in addition needed all the additional intense, superior processors that gave the tools to reveal the bits of data in huge information. Today's IoT-associated systems predominate the information volume from this play of huge information. As gadgets & sensors keep multiplying, thus can the amount of information they create.

For instance, a solitary motor vehicle enormous auto can turn out 4000 GB of information for every day. The new airliner A380-1000 is equipped 10000 sensors in every wing. Inheritance huge information innovation will not be able to handle the information created by associated apparatuses in shrewd homes, movement sensors in sensible urban areas, & automatic systems in savvy factories.

5.2. New & exciting system requirements

ML is important to breaking down the large, boring volumes of information spilling out of tremendous, faithfully on IoT systems. whereas ML could appear as if sci-fi to varied, it's as of currently getting used & natural to purchasers of net-based social networking & web primarily based looking (Facebook's news encourage depends on ML calculations, & Amazon's proposal motor uses ML to advocate what book or movie you got to appreciate next).

ML systems understand the standard stream samples of information show on IoT systems & focus on the peculiarities or examples outside the quality. Thus from billions of information focuses, ML will isolate the "flag from the clamour" in large information flows, serving to associations focus on what is necessary. Regardless, to be useful & fruitful for associations, ML computations must run counts at monster scale in a matter of milliseconds — on a constant introduce. These constantly complex counts put weight on traditional data-center processors & figuring stages.

To work at scale & continuously, ML frameworks require processors with various joined focuses, faster memory subsystems, & plans that can parallelize getting ready for front line consistent understanding. These are stages with worked in illustrative taking care of engines & the capacity to run complex computations in-memory for consistent results & snappy utilization of bits of learning.

5.3. Final prediction

Processors worked for predominant enrolling will be looked for after. ML & fake awareness will require fundamentally more power as they arrive at an undeniable conclusion with respect to IoT information streams & client engagement for improved arrangements & exertion.

These processors were by & large the locale of research labs & supercomputing challenges, for instance, exhibiting atmosphere cases & genome sequencing. In any case, ML stages will turn out to be progressively essential as IoT frameworks end up noticeably greater & more certain — & as associations dynamically build their accomplishment in light of the bits of information found in machine-to-machine correspondence. These processors pass on the execution required for the most requesting workloads, including ML & fake cognizance figurines. So, they will never again be restricted to the rarified states of supercomputing in examine centers & universities, as they continuously turn into an essential for cutting edge associations.

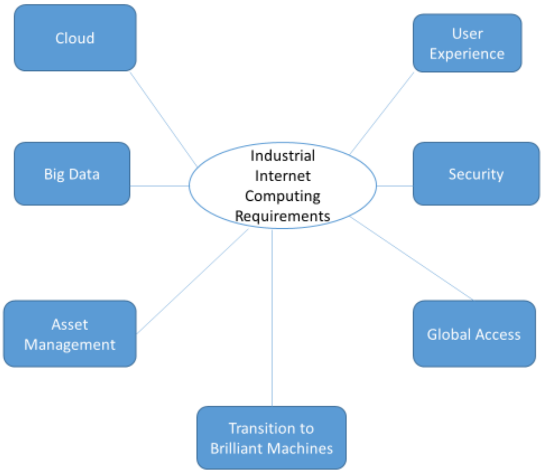


Fig. 3. IoT's computing requirements

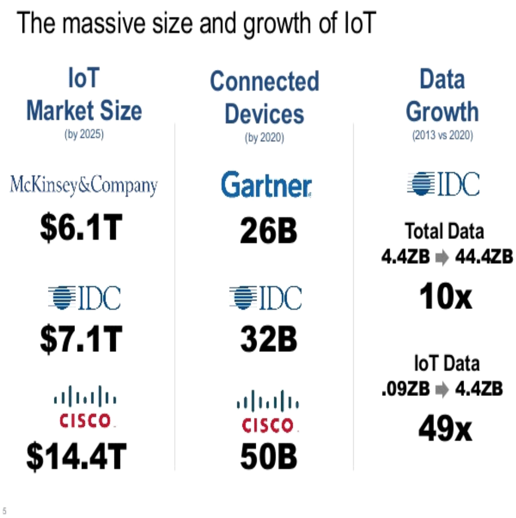


Fig. 4. Size growth of IoT

6. Conclusion

This paper endeavoured to give further knowledge about the IoT & BD structures, the impacts of IoT on BD, the BD propels & the troubles. Since there is a vital impact of IoT on BD we need to quickly improvise the aggregate structure to manage the step by step developing conditions. There are two or three districts of concern & security & insurance & data amassing profitability are likely the most troublesome issues we are defying. Security exchange off & inefficient perspectives in data gathering frameworks result in lost status, money, time & effort. Regardless, there is trust in light of the fact that both the IoT & the BD are at a creating stage & there will be update.

References

[1]. Sowe, S. K., Kimata, T., Dong, M., & Zettsu, K. (2014, July). Managing heterogeneous sensor data on a big data platform: IoT services for data-intensive science. In Computer Software and Applications Conference Workshops (COMPSACW), 2014 IEEE 38th International (pp. 295-300). IEEE.https://doi.org/10.1109/COMPSACW.2014.52

[2]. Sivaraman, V., Gharakheili, H. H., Vishwanath, A., Boreli, R., & Mehani, O. (2015, October). Network-level security and privacy control for smart-home IoT devices. In Wireless and Mobile Computing, Networking and Communications (WiMob), 2015 IEEE 11th International Conference on (pp. 163-167). IEEE.https://doi.org/10.1109/WiMOB.2015.7347956

[3]. Kaul, L., & Goudar, R. H. (2016, November). Internet of things and Big Data-challenges. In Green Engineering and Technologies (IC-GET), 2016 Online International Conference on (pp. 1-5). IEEE.https://doi.org/10.1109/GET.2016.7916854

[4]. Perera, C., Ranjan, R., Wang, L., Khan, S. U., & Zomaya, A. Y. (2015). Big data privacy in the internet of things era. IT Professional, 17(3), 32-39.https://doi.org/10.1109/MITP.2015.34

[5]. Kumar, J. S., & Patel, D. R. (2014). A survey on internet of things: Security and privacy issues. International Journal of Computer Applications, 90(11).

[6]. Gudivada, V. N., Baeza-Yates, R. A., & Raghavan, V. V. (2015). Big Data: Promises and Problems. IEEE Computer, 48(3), 20-23.

[7]. Hwang, Y. H. (2015, April). Iot security & privacy: threats and challenges. In Proceedings of the 1st ACM Workshop on IoT Privacy, Trust, and Security (pp. 1-1). ACM.https://doi.org/10.1145/2732209.2732216

[8]. Cook, D. J., Thompson, J. E., Prinsen, S. K., Dearani, J. A., & Deschamps, C. (2013). Functional recovery in the elderly after major surgery: assessment of mobility recovery using wireless technology. The Annals of thoracic surgery, 96(3), 1057-1061.http://dx.doi.org/10.1016/j.athoracsur.2013.06.057

[9]. Etherington, D. (2014). Amazon Echo Is A $199 Connected Speaker Packing an Always-On Siri-Style Assistant. Tech Crunch, 6.

[10]. M. Mendoza. (2014). Jaguar Land Rover Develops Self-Learning, Intelligent Car. Tech Times. 17 July 2014

[11]. Tim Allen. (2017). How to Solve IoT's Big Data Challenge with Machine Learning. Digitalist Magazine. Internet of Things.

[12]. Devi, T.J.B., Subramani, A., & Solanki, V. (2017). Smart City: IOT Based Prototype for Parking Monitoring and Management System Commanded by Mobile App. Volume 10, ACIS, Poland. DOI 10.15439/2017R23

[13]. Dhall & Solanki. (n.d.). An IoT Based Predictive Connected Car Maintenance Approach. International Journal of Interactive Multimedia and Artificial Intelligence. Volume 04, No 03, DOI 10.9781/ijimai.2017.433

[14]. Son, L.H., Jha, S., Kumar, R. et al. Telecommun Syst (2018). https://doi.org/10.1007/s11235-018-0481-x

Author’s Biography

Jyotir Moy Chatterjee: Jyotir Moy Chatterjee worked as Assistant Professor in Department of Computer Science and Engineering at GD-RCET, Bhilai, C.G, India. He received M. Tech from KIIT University, Bhubaneswar, Odisha and B. Tech in Computer Science & Engineering from Dr. MGR Educational & Research Institute University. He is the member of CSI. His research interests include the cloud computing, big data, privacy preservation and data mining. He is also Oracle Certified OCA 10g and IBM Certified Associate System Administrator Lotus Notes and Domino 8. He is the member of IAENG, IRED, IAIP, ICSES. He has more than 25 International Journal publication. He Served as Co-convener at International Multidisciplinary Conference on Emerging Trends in Engineering, Science and Technology 2017 at Sanjay Rungta Group of Institutions, Bhilai. He reviewed research papers for ACM Computing Surveys Journal (SCI indexed). He was selected as Subject Expert in Sodh, Samiksha Aur Mulyankan Journal (UGC approved). He got selected as reviewer of Oriental Journal of Computer Science and Technology. He also was selected as a reviewer for IJECE(Scopus). He received best paper award for "Fog Computing-Beginning of a New Era" in PARAS 2017 (Parthivi College of Engineering & Management, Bhilai).

How to Cite

Chatterjee, Jyotir Moy, "IoT with Big Data Framework using Machine Learning Approach*",International Journal of Machine Learning and Networked Collaborative Engineering,* Vol. 02 No. 02, 2018, pp. 075-085.
doi: https://doi.org/10.30991/IJMLNCE.2018v02i02.005.