

# Zerocrossing Intervals And Distribution Using SNR and ANN

Rirtan Singh kumar, Manish Rathod

Assistant Professor, Dept of EEE, Indus University, India

## Article Information

Received : 02 Sept 2024  
Revised : 06 Sept 2024  
Accepted : 18 Sept 2024  
Published : 22 Sept 2024

## Corresponding Author:

Rirtan Singh kumar

**Abstract**— Speech signal is modelled using the average energy of the signal in the zerocrossing intervals. Variation of these energies in the zerocrossing interval of the signal is studied and the distribution of this parameter through out the signal is evaluated. It is observed that the distribution patterns are similar for repeated utterances of the same vowels and varies from vowel to vowel. Credibility of the proposed parameter is verified over five Malayalam (one of the most popular Indian language) vowels using multilayer feed forward artificial neural network based recognition system. The performance of the system using additive white Gaussian noise corrupted speech is also studied for different SNR levels. From the experimental results it is evident that the average energy information in the zerocrossing intervals and its distributions can be effectively utilised for vowel phone classification and recognition.

**Keywords:** *Zerocrossing Intervals, Phoneme Recognition, ANN*

**Copyright © 2024: Rirtan Singh kumar**, This is an open access distribution, and reproduction in any medium, provided Access article distributed under the Creative Commons Attribution License the original work is properly cited License, which permits unrestricted use.

**Citation:** Rirtan Singh kumar, “Zero crossing Intervals And Distribution Using SNR and ANN”, Journal of Science, Computing and Engineering Research, 7(9), September 2024.

## I. INTRODUCTION

Parameterization of analog speech signal is the first step in the speech recognition process. Although various aspects of phoneme recognition have been investigated by researchers, the parameterization of analog speech and its computational complexity is still a headache for them. We are interested in a set of processing techniques used for phone recognition that are reasonably termed as time domain methods. By this we mean that the processing methods involve the wave form of the speech signal directly. Some examples of representation of speech signal in terms of time domain measurements include average zerocrossing rate, energy and autocorrelation function. The time domain method like zerocrossing analysis technique has been applied to several signal processing and signal analysis tasks. Some of this task include speech analysis and speech recognition [1] [2] [3] [4], electroencephalographic (EEG), biomedical applications, communication application, oceanographic analysis and many others [5]. The relative easy method by which Zerocrossing information can be extracted and its low cost implementation made this technique attractive. Zerocrossing rate is widely used in many speech analysis and recognition purposes. In application involving speech processing and speech recognition, additional interest in Zerocrossing analysis has gained impetus through observation of Licklider and Pollack who showed that clipped speech is highly intelligible [6]. As a result, numerous speech analysis and recognition devices have been built utilizing Zerocrossing analysis techniques [7]. In spite of this simplicity, the resulting representation provides a useful basis for estimating important features of the speech signal. Phoneme recognition has major

contribution in designing practical efficient speech recognition systems. Many phoneme recognition systems are reported in literature [8] [9] [10] [11]. In this paper we modelled the speech signal with a computationally simple speech parameter using the average energy information in the zerocrossing intervals of the signal and used it for phoneme recognition applications. We have used five Malayalam (one of the most popular Indian language) vowels for the conduct of recognition experiments. The present work also illustrates how the proposed parameters can be effectively used in neural network based phoneme recognition systems. The paper is organized in five sections. Session 1.1 describes the representation of energy in the discrete speech signals. Session 2 describes the speech modelling technique using average energy in the zerocrossing interval and the Session 3 describes the speech parameterization using the average energy in the zerocrossing interval distribution. Session 4 describes vowel recognition using Artificial Neural Network (ANN) and the final section deals the conclusion.

1.1. Energy of the Discrete Speech Signal In a typical speech signal we can see that its certain properties considerably changes with time. For example, we can observe a significant variation in the peak amplitude of the signal and a considerable variation of fundamental frequency within voiced regions in a speech signal. These facts suggest that simple time domain processing techniques should be capable of providing useful information of signal features, such as intensity, excitation mode, pitch, and possibly even vocal tract parameters, such as formant frequencies. Most of the short time processing techniques that give time domain features (Qn), can be mathematically

represented as  $\infty = -\infty$  - where  $T[]$  is the transformation matrix which may be either linear or nonlinear,  $X(m)$  represents the data sequence and  $W(n-m)$  represents a limited time window sequence. The energy of the discrete time signal is defined as  $E_n = \sum_{m=0}^{n-1} |X(m)|^2$ . Such a quantity has little meaning or utility for speech since it gives little information about time dependent properties of speech signal. We have observed that the amplitude of the speech signal varies appreciably with time. In particular, the amplitude of the unvoiced segment is generally much lower than amplitude of the voiced segment. The short time energy of the speech signal provides a convenient representation that reflects the amplitude variation and can be defined as  $E_n = \sum_{m=0}^{n-1} |X(m)|^2$ . The major significance of  $E_n$  is that it provides a basis for distinguishing voiced speech segment from unvoiced speech segment. It can be seen that the value of  $E_n$  for the unvoiced segments are significantly smaller than voiced segments. The energy function can also be used to locate approximately the time at which voiced speech become unvoiced speech and vice versa, and for high quality speech (high signal to noise ratio) the energy can be used to distinguish speech from silence. The above discussion cites the importance of energy function ( $E_n$ ) for speech analysis purpose.

## II. RELATED WORKS SPEECH MODELLING USING AVERAGE ENERGY IN THE ZEROCROSSING

The speech production model [12] suggests that the energy of the voiced speech is concentrated below about 3 kHz, where as in the case of unvoiced speech, most of the energy is found at higher frequencies. Since high frequency implies high zerocrossing rate and low frequency implies low zerocrossing rate, there is strong correlation between zerocrossing rate and energy distribution with frequency [13] [14]. This motivates us to model the speech signal using average energy in zerocrossing interval of the signal. Consider the speech segment shown in Figure 1. The  $ZC^k_i$  shows the  $i$ th zerocrossing and  $ZC^k_{i+1}$  shows the  $i+1$ th zerocrossing of  $k$ th observation window. The time interval between these two points is called  $i$ th zerocrossing interval  $T^k_i$  in the  $k$ th observation window

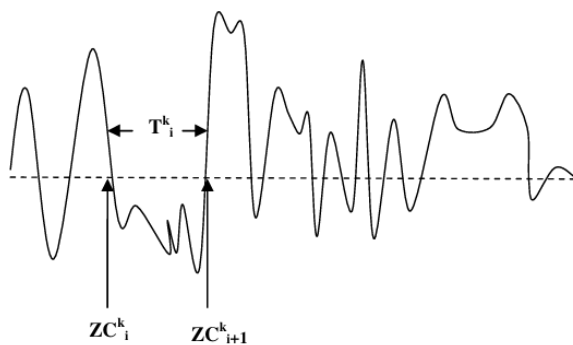


Figure 1. Speech segment in  $k^{th}$  observation window.

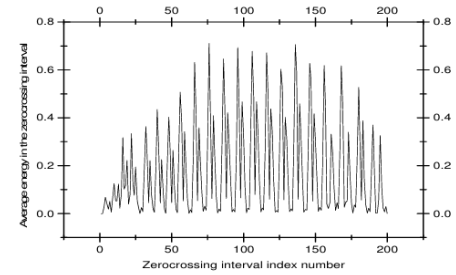


Figure 2 (a) Average energy in the zerocrossing interval vs zerocrossing interval index number of vowel /a/

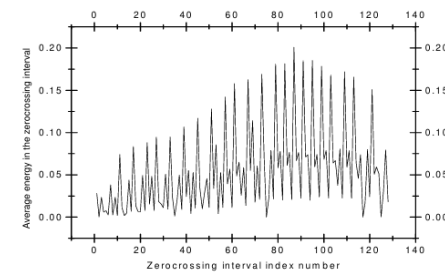


Figure 2 (b) Average energy in the zerocrossing interval vs zerocrossing interval index number of vowel /i/

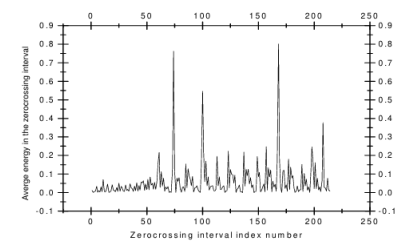


Figure 2 (d) Average energy in the zerocrossing interval vs zerocrossing interval index number of vowel /e/

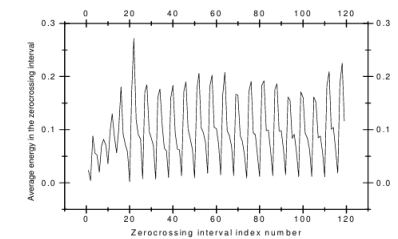


Figure 2 (e) Average energy in the zerocrossing interval vs zerocrossing interval index number of vowel /o/

The average energy in the  $i$ th zerocrossing interval can be obtained by the expression  $E_{i,k} = \frac{1}{T^k_i} \int_{ZC^k_i}^{ZC^k_{i+1}} |X(t)|^2 dt$  is the average energy of the signal in  $T^k_i$  th zerocrossing interval of  $k$ th observation window and  $X(t)$  is the instantaneous signal amplitude. The aim of the present study is to find a robust coefficient for speech recognition application using the average energy in the zerocrossing interval (AEZI). An XY plot is generated by plotting index number of zero crossing interval along X axis and Average Energy in the

Zerocrossing Interval (AEZI) along Y axis. Figure 2 (a-e) represents the average energy in the zerocrossing interval vs index number of the zerocrossing interval for the Malayalam vowels /a/, /i/, /u/, /e/, /o/ respectively.

### III. SPEECH MODELLING USING AVERAGE ENERGY IN THE ZEROCROSSING

Speech signal is low pass filtered to 4kHz and sampled at 8kHz rate and digitized using 16 bit A/D converter. Five Malayalam vowels /a/, /i/, /u/, /e/, and /o/ uttered by a single speaker is used in the present study. From the band limited speech data, we have extracted Average Energy in the Zerocrossing Intervals (AEZI) using the equation =  $1/2 + 1$ . For finding the distribution of average energy in the zerocrossing interval for each vowel, first we fix a maximum threshold of energy value as 0.8 by examining the computed energy values of all the vowels. This upper threshold is divided into twenty uniform ranges (0-0.04, 0.04-0.08, ..., 0.76-0.8). The distribution of the average energy of the vowels in the zerocrossing interval among all the ranges mentioned above is obtained next. Figure 3(a-e) shows the distribution of average energy in the zerocrossing intervals of Malayalam vowels /a/, /i/, /u/, /e/ and /o/. The normal slim lines represent the distribution graph obtained using the repeated utterances of same vowels by the same speaker. The bold line shows the mean distribution plot.

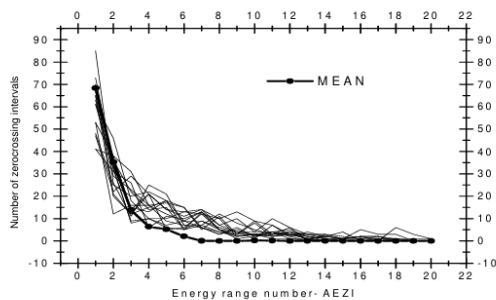
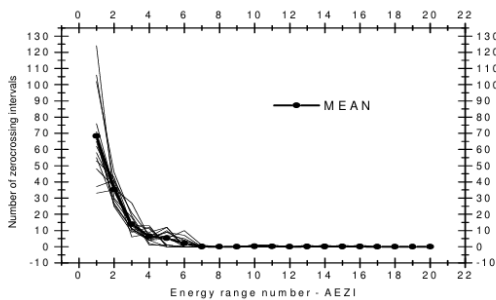


Figure 3 (a) Number of zerocrossing intervals vs Energy range number –AEZI of the vowel /a/



The application of artificial neural networks to speech recognition is the youngest and least understood of the recognition technologies. The ANN is based on the notion

that complex “computing” operations can be implemented by massive integration of individual computing units, each of which performs an elementary computation. Artificial neural networks have several advantages relative to sequential machines. First, the ability to adapt is at the very center of ANN operations. Adaptation takes the form of adjusting the connection weights in order to achieve desired mappings. Furthermore ANN can continue to adapt and learn, which is extremely useful in processing and recognition of speech. Second, ANN tend to move robust or fault tolerant than Von Neumann machines because the network is composed of many interconnected neurons, all computing in parallel, and failure of a few processing units can often be compensated for by the redundancy in the network. Similarly ANN can often generalize from incomplete or noisy data. Finally ANN when used as classifier does not require strong statistical characterization or parameterization of data [15] [16] [17] [18] [19]. These are the main motivations to choose artificial neural networks for phoneme recognition. We used Multilayer Layer Feed Forward architecture for this experiment. The network consists of 20 input nodes and five output nodes, representing the five vowels. The network is trained using the energy information of the signal mentioned in Section 3. The experiment is repeated by changing the number of hidden layers and adding Additive White Gaussian Noise (AWGN) to the signal with different Signal to Noise Ratio (SNR). We used the database of 150 samples of each vowel spoken by a single male speaker for this purpose. A disjoint set of training and test patterns are formed by taking the 50% of the total patterns for each set. The error tolerance (Emax) is fixed as 0.001 and learning parameter ( $\eta$ ) is chosen between 0.01 and 1.0. The training of the network is done using error back propagation learning method. The recognition results are tabulated in Table 1 for neural network with three hidden layers. The recognition results indicates that the network shows poor recognition accuracy when the signal is added with Additive White Gaussian Noise of 0dB, 3dB and 10dB SNR. Above 20dB SNR, it shows better results. The average vowel recognition accuracy obtained for five Malayalam vowels in this experiment is 87.94%. This percentage is comparable with the phoneme recognition using spectral method based on ANN.

### IV. METHODOLOGY

Table 1. Recognition accuracy for five Malayalam vowels with additive white Gaussian noise of different dB levels.

Vowel	Recognition accuracy %					
	0dB	3dB	10dB	20dB	30dB	Normal
/a/	33.3	33.3	46.6	66.6	86.6	86.6
/i/	26.6	33.3	40.0	53.3	86.6	86.6
/u/	13.3	26.6	33.3	73.3	86.6	86.6
/e/	13.3	33.3	40.0	60.0	93.3	93.3
/o/	26.6	33.3	46.6	73.3	86.6	86.6
Average	22.62	31.96	41.3	65.3	87.94	87.94

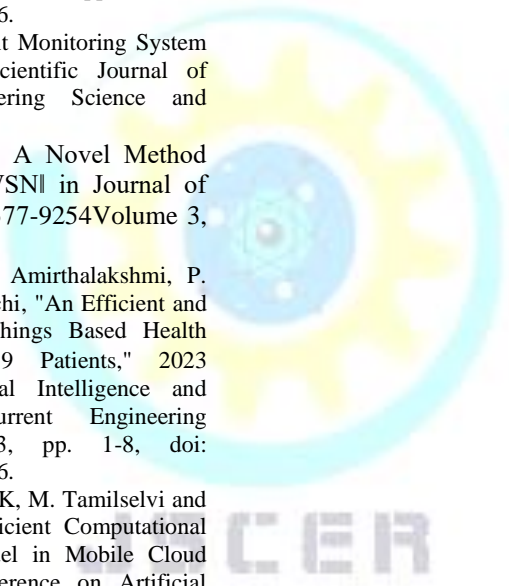


## V. CONCLUSION

The speech signal is modelled using the average energy in the zerocrossing interval of the signal. Variation of these energies in the zerocrossing interval of the signal is studied and the distribution of this parameters through out the signal is evaluated. We found that the distribution patterns are almost similar for repeated utterances of the same vowels and varies from vowel to vowel. This distribution pattern is used for recognizing five Malayalam vowels using multilayer feed forward artificial neural network. The performance of this system using additive white Gaussian noise corrupted speech (vowel) is also studied for different SNR levels. It is found that recognition accuracy of the vowels is good when the signal is corrupted with low noise levels. The recognition accuracy obtained for five Malayalam vowels is 87.94 %. The present work can be extended, with the proposed neural network model and training algorithm, using vowels samples uttered by different male and female speakers under different age groups for analyzing the speaker independency of the proposed parameters.

## REFERENCES

- [1]. P. Nirmala, T. Manimegalai, J. R. Arunkumar, S. Vimala, G. Vinoth Rajkumar, Raja Raju, "A Mechanism for Detecting the Intruder in the Network through a Stacking Dilated CNN Model", *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 1955009, 13 pages, 2022. <https://doi.org/10.1155/2022/1955009>.
- [2]. D. Sathyanarayanan, T. S. Reddy, A. Sathish, P. Geetha, J. R. Arunkumar and S. P. K. Deepak, "American Sign Language Recognition System for Numerical and Alphabets," 2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE), Chennai, India, 2023, pp. 1-6, doi: 10.1109/RMKMATE59243.2023.10369455.
- [3]. J. R. Arunkumar, Tagele berihun Mengist, 2020" Developing Ethiopian Yirgacheffe Coffee Grading Model using a Deep Learning Classifier" *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Volume-9 Issue-4, February 2020. DOI: 10.35940/ijitee.D1823.029420.
- [4]. Ashwini, S., Arunkumar, J.R., Prabu, R.T. et al. Diagnosis and multi-classification of lung diseases in CXR images using optimized deep convolutional neural network. *Soft Comput* (2023). <https://doi.org/10.1007/s00500-023-09480-3>
- [5]. J.R.Arunkumar, Dr.E.Muthukumar," A Novel Method to Improve AODV Protocol for WSN" in *Journal of Engineering Sciences*" ISSN NO: 0377-9254 Volume 3, Issue 1, Jul 2012.
- [6]. R. K, A. Shameem, P. Biswas, B. T. Geetha, J. R. Arunkumar and P. K. Lakineni, "Supply Chain Management Using Blockchain: Opportunities, Challenges, and Future Directions," 2023 Second International Conference on Informatics (ICI), Noida, India, 2023, pp. 1-6, doi: 10.1109/ICI60088.2023.10421633.
- [7]. Arunkumar, J. R. "Study Analysis of Cloud Security Chanllenges and Issues in Cloud Computing Technologies." *Journal of Science, Computing and Engineering Research* 6.8 (2023): 06-10.
- [8]. J. R. Arunkumar, R. Raman, S. Sivakumar and R. Pavithra, "Wearable Devices for Patient Monitoring System using IoT," 2023 8th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2023, pp. 381-385, doi: 10.1109/ICCES57224.2023.10192741.
- [9]. S. Sugumaran, C. Geetha, S. S, P. C. Bharath Kumar, T. D. Subha and J. R. Arunkumar, "Energy Efficient Routing Algorithm with Mobile Sink Assistance in Wireless Sensor Networks," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ACCAI58221.2023.10201142.
- [10]. R. S. Vignesh, V. Chinnammal, Gururaj.D, A. K. Kumar, K. V. Karthikeyan and J. R. Arunkumar, "Secured Data Access and Control Abilities Management over Cloud Environment using Novel Cryptographic Principles," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-8, doi: 10.1109/ACCAI58221.2023.10199616.
- [11]. Syamala, M., Anusuya, R., Sonkar, S.K. et al. Big data analytics for dynamic network slicing in 5G and beyond with dynamic user preferences. *Opt Quant Electron* 56, 61 (2024). <https://doi.org/10.1007/s11082-023-05663-2>
- [12]. Krishna Veni, S. R., and R. Anusuya. "Design and Study Analysis Automated Recognition system of Fake Currency Notes." *Journal of Science, Computing and Engineering Research* 6.6 (2023): 16-20.
- [13]. V. RamKumar, S. Shanthi, K. S. Kumar, S. Kanageswari, S. Mahalakshmi and R. Anusuya, "Internet of Things Assisted Remote Health and Safety Monitoring Scheme Using Intelligent Sensors," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-8, doi: 10.1109/ACCAI58221.2023.10199766.
- [14]. R. S. Vignesh, R. Sankar, A. Balaji, K. S. Kumar, V. Sharmila Bhargavi and R. Anusuya, "IoT Assisted Drunk and Drive People Identification to Avoid Accidents and Ensure Road Safety Measures," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ACCAI58221.2023.10200809.
- [15]. I. Chandra, G. Sowmiya, G. Charulatha, S. D, S. Gomathi and R. Anusuya, "An efficient Intelligent Systems for Low-Power Consumption Zigbee-Based Wearable Device for Voice Data Transmission," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICECONF57129.2023.10083856.
- [16]. G. Karthikeyan, D. T. G, R. Anusuya, K. K. G, J. T and R. T. Prabu, "Real-Time Sidewalk Crack Identification and Classification based on Convolutional Neural Network using Thermal Images," 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, 2022, pp. 1266-1274, doi: 10.1109/ICACRS55517.2022.10029202.

- 
- [17].R. Meena, T. Kavitha, A. K. S, D. M. Mathew, R. Anusuya and G. Karthik, "Extracting Behavioral Characteristics of College Students Using Data Mining on Big Data," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICECONF57129.2023.10084276.
- [18].S. Bharathi, A. Balaji, D. Irene. J, C. Kalaivanan and R. Anusuya, "An Efficient Liver Disease Prediction based on Deep Convolutional Neural Network using Biopsy Images," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2022, pp. 1141-1147, doi: 10.1109/ICOSEC54921.2022.9951870.
- [19].I. Chandra, G. Sowmiya, G. Charulatha, S. D, S. Gomathi and R. Anusuya, "An efficient Intelligent Systems for Low-Power Consumption Zigbee-Based Wearable Device for Voice Data Transmission," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICECONF57129.2023.10083856.
- [20].Revathi, S., et al. "Developing an Infant Monitoring System using IoT (INMOS)." International Scientific Journal of Contemporary Research in Engineering Science and Management 6.1 (2021): 111-115.
- [21].J.R.Arunkumar, Dr.E.Muthukumar, A Novel Method to Improve AODV Protocol for WSN in Journal of Engineering Sciences ISSN NO: 0377-9254 Volume 3, Issue 1, Jul 2012.
- [22].R. S. Vignesh, A. Kumar S, T. M. Amirthalakshmi, P. Delphy, J. R. Arunkumar and S. Kamatchi, "An Efficient and Intelligent Systems for Internet of Things Based Health Observance System for Covid 19 Patients," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-8, doi: 10.1109/ICECONF57129.2023.10084066.
- [23].I. Chandra, K. V. Karthikeyan, R. V, S. K, M. Tamilselvi and J. R. Arunkumar, "A Robust and Efficient Computational Offloading and Task Scheduling Model in Mobile Cloud Computing," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-8, doi: 10.1109/ICECONF57129.2023.10084293.
- [24].R. K, A. Shameem, P. Biswas, B. T. Geetha, J. R. Arunkumar and P. K. Lakineni, "Supply Chain Management Using Blockchain: Opportunities, Challenges, and Future Directions," 2023 Second International Conference on Informatics (ICI), Noida, India, 2023, pp. 1-6, doi: 10.1109/ICI60088.2023.10421633.
- [25].J. R. Arunkumar, and R. Anusuya, "OCHRE: A Methodology for the Deployment of Sensor Networks." American Journal of Computing Research Repository, vol. 3, no. 1 (2015): 5-8.