ZIKA Virus: A Secure System USINF NBN Classifier for Predicting and Preventing ZIKA in Cloud

B. Mahalakshmi, G. Suseendran

Abstract: ***One of the upcoming mosquitoes borne disease is Zika Virus which is spreading rapidly around the world. The traditional ways of detecting Zika Virus are not much effective. To overcome a proposed method of cloud based system is developed with the combination of cloud computing mobile phones, fog computing finally IoT for sensing the mosquitoes. Here fog computing acts as a bond between cloud and user for reducing the latency level and increasing the processing speed. A NBN- Naïve Bayesian Network algorithm is to find the infected patient and uninfected one, mosquitoes breeding sites and dense site by mapping it in GPS and finding the location. The cloud based system provides high accuracy in predicting the result using NBN classifier and also shows the risk prone site location to the government health sector. It provides the user a better communication with the health sector, to avoid the outbreak of ZVD and, the complete cloud system gives better high accuracy for prediction and preventing.***

Keywords: Zika Virus, Fog Computing, NBN classifier, Cloud Computing, Data protection.

I. INTRODUCTION

Zika virus is another mosquito borne disease which threatens the human life. It was a virus found in sentinel rhesus monkey from zika forest in Uganda in 1947[1]. Aedes mosquitoes are found in African forest which is isolated. It was found in human in the 1952 in Uganda and The United Republic of Tanzania. According to World Health Organization (WHO) since fro 2007 to 2016 around sixty four countries and territories has been detected as infected [2]. In India the recent report from Quartz India shows that a recent outbreak of Zika is found again in less than two years. It has been detected in Jaipur, around 22 people got infected through the virus. And one person from Bihar is on alert. It was first outbreak in Gujarat and in Tamilnadu in the year 2017. In 2018 Jaipur is found to be infected. Aedes mosquito was the reason for Zika, incidentally Dengue, and Chikungunya is the vector cause from Aedes mosquitoes.

The ZVD symptoms are same as Dengue, and Chikungunya, which consist of fever, skin rashes, conjunctivitis, joint pain, muscle pain, headache and exposure to risk area. The major issue of ZVD is it leads to death, apart from that once its affect the pregnant women it directly affects the newborn babies with the disease called microcephaly and some neurological disorder. Other than mosquitoes ZVD transmitted through sexual transmission, prenatal transmission from mother to fetus during their pregnancy, sexual intercourse. ZVD spreads rapidly in India and all over the world. As of now there is no vaccine for ZVD and no medicine for treatment, the only way is to prevent ourselves from mosquito bites and should avoid the area of viral infected. The ZVD symptoms are generally mild so it is a great risk to carrying mother and the fetus [3]. Figure 1 represents the details about Zika Virus, their symptoms and causes.

Figure 1: Zika Virus Symptoms, causes

ZVD becomes a great challenge to government to provide better healthcare services. The communication between the healthcare centre and the people is the only way to stop the fast spreading ZVD. Nowadays there is great opportunity to find the infection and the outbreak of disease from the remote location itself. The current technologies have collaboration among themselves to give quality computing service. The advancement in the network technology with cloud services, GIS-Geographic Information System and smart phone technologies will provide better service to the people. There is an on demand service provided in the technology. So many healthcare services are moved to the current technologies, which we get a real time services remotely. A fog based cloud approach is used for monitoring the patient and detects the ZVD through some effective methods.

II. RELATED WORKS

In the year 1952 first ZVD was found in human, later it was found in many countries in Asia and Africa, then it is spread to the Pacific region and lately to America the outbreak there is rapid spreading of Zika virus globally.

Paixao et al.[4] examined the symptoms and tricky situation developing from zika virus.

Revised Manuscript Received on December 30, 2018.

B. Mahalakshmi, Ph.D. Research Scholar, Department of Computer Science, School of Computing Science, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, Tamil Nadu, India.

G. Suseendran, Assistant Professor, Department of Information Technology, School of Computing Science, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, Tamil Nadu, India.

The ZVD outbreak found in Brazil first where 14,835 cases of infected were identified. From the infected ZVD cases it was found that around 2.3 per 1000 had neurological disease and 1.3 per 1000 cases found Guillain-Barre syndrome. The ZVD outbreak the microcephaly disease found in the fetus of the pregnant mother. Nishiura et al.[5] has reviewed, in Brazil year 2015 the outbreak is occurred and shows about the risk level of ZVD among pregnant women and the risk of microcephaly too. Petersen et al.[6] the author studied about the impact of ZVD outbreak in America, which in need of the public health. It also reviewed the symptoms and outbreak of ZVD in Brazil and insisting the urgent for protecting and prevention of the spread of ZVD.

Lopez-Barbosa et al. [7] has proposed about the real time use of IOT devices, smart phones. Senor devices, cloud computing, and explain how it is useful for the patient analysis and getting the accurate result through real time.

Quwaider and Jararweh [8] has proposed about the public health related awareness using the sensor devices and cloud computing. Map reduce concept is used to find the abnormalities in the information produced real time values by the sensors. Mamun et al. [9] has explained a concept of transmitting the voice signals to the doctor through cloud technology. The doctor then diagnosis the patient and monitoring them with the help of mobile phones and cloud is used for the proposed methodology. Zhang et al. [10] proposed a monitoring and controlling the outbreak using smart phone technologies. Here the overall population is splitted as numerous clusters and the outbreak methods are used and applied in the cluster level based on the network contact. Sareen, Sood, and Gupta[11] has proposed about the invasive technological improvement in IoT, Mobile and cloud computing. By improving the quality of service of the technology healthcare services are improved simultaneously.

III. PROPOSED SYSTEM

The proposed system contains four different modules. The data accretion, fog layer, cloud layer and finally process is communicating with the healthcare related individuals. A framework is structured for finding the ZVD and also for data protection. Fog computing is used here as it plays a major role as it can process the big data which are collected from the sensor or mobile devices. It is a bridge for the end user and large scale cloud services for storing, Computing and application services to the data. Figure 2 shows the detailed architecture of the proposed model for predicting the ZVD.

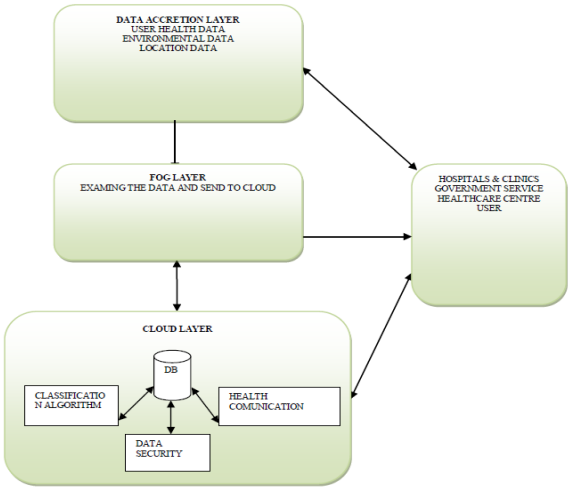

Figure 2: Architecture of the Proposed Model for predicting ZVD

Data Accretion Layer

In the proposed model data module contains the details of the user health data, environmental data and location data. Every user has to register in the system through the application provided in the mobile. Each user got a record number (Rno) created by the system. The symptoms of Zika Virus are gathered once in a while from the users. The symptoms are answered in the yes or no pattern. Not only the symptoms the user health related records also gathered. In the health data the detailed symptoms of patient for instance fever temperature, body

Pain, joint pain, headaches etc are collected. These types of data are collected with the help of the sensor provided to the user. The collected data are secured using some encryption method in a priority. [12]

The symptoms of ZVD are gathered over a period of time and the details are entered as yes or no pattern by the user. The environmental sensor collects the data about the breeding of mosquito and the density of it. The sensor captures the data continuously and finds the location of the breeding site. And also the carbon dioxide levels, temperature of air and humidity values are frequently analyzed so that the climatic condition of the particular location is identified. Every data related to environment are collected in this part and stored in the fog computing servers. The location section comprises of the data which are related to previous one as it shows the perfect location where there is a possibility of mosquito density and breeding chances are high comparing with the climatic conditions. Table 1(a), (b), (c) shows the attributes of the proposed model.

Table 1(a): Personal attributes

S.NO	Attributes	Narration
1	Ref no	Reference no of the user
2	Name	Name of the user
3	Age	Age of the user
4	Gender	Male/Female
5	Address	Address of the user
6	Mobile No	Mobile no of the user

Table 1(b): Symptoms of ZVD

S.NO	Attributes	Answer
1	Fever	Y/N
2	Joint pain	Y/N
3	Conjunctivitis	Y/N
4	Skin Rashes	Y/N
5	Muscle pain	Y/N
6	Head ache	Y/N
7	Exposure to risk area	Y/N

Table 1(C): Environmental attributes

S.NO	Attributes	Narration
1	Mosquito dense area	Density of mosquito in GPS location
2	Mosquito breeding area	Breeding location in GPS
3	Temperature	Temperature of the stagnant water
4	Humidity	Humidity level
5	Carbon dioxide	Co_2 value

Fog Layer

Fog computing is a decentralized environment which is used to process big data values for authentic time. It is a bridge between the user and the cloud computing where the large storage of data is possible in cloud. In order to prevent the processing and performance time the fog computing is used[22]. It collects all the sensor data and stored it in the fog server, only the required data is analyzed and send to cloud for the data processing. So that the processing speed and time are saved.

The latency range bandwidth everything are increased because of fog. So it is independent server for data processing and storing. In the proposed work fog having the job of collecting all sensitive info from user and examine whether the symptoms are matching with the user. So that type of result only sends to the cloud. Fog is a first level of environment where the sensitive data collected from the sensor are needed to be stored in a large number. So there is a need to examine those data and should match with the specified one. Next it sends those data into cloud where the final data classification and further process are going to get done.

Data Security

The collected data are protected by using the secret sharing algorithm where the data are fragmented into small and prioritization is provided for the different levels[13]. The security is given on the basis of sensitivity of the data. The higher prioritization is given to the user personal data where it should be protected from the unauthorized person [16]. The second level is the environmental data where the data are stored in different server [14]. The third is least as it should contain the data of symptoms and prior notice to the person to take necessary steps for avoiding the ZVD. The government health sectors hospitals are giving the necessary advice to the user.

3.4. Classification- NBN Based Classifier

This module provides the user basic analysis and verdict by classifying the ZVD parameters values as Infected and Uninfected using NBN classifier and categorized as IN (Infected) and UIN (Uninfected). This classifier is probabilistically strong for classification problems in multiple domains. In this method all variables $A_1,A_2,A_3\ldots..A_n$ are independently given a class variable C. Because the ZVD symptoms are independent of each single person. Using NBN the user is classified into IN(infected) or UIN(Uninfected). The probability attributes are comes from the training data set for proper estimation. Consider C as a class variable that represents the categories as IN(infected) and UIN(uninfected) based on a vector symptoms S_i=(Fv, Skr, Cj,Jtp,Msp,Ha,Ra). For classifying a user $a_i \in$ A=1,2,….n as infected or uninfected using the symptoms of the individual user S_i, using the Baye's rule, probability of every class is calculated is given below:

$$P(C = UIN|Si) = \frac{P(UIN)P(Si|UIN)}{P(Si)}$$

$$P(C = IN|Si) = \frac{P(IN)P(Si|IN)}{P(Si)}$$

Where $P(C = UIN|Si)$ tells about the probability of user who is uninfected (UIN) based on the ZVD S_i symptoms. $P(C = IN|Si)$shows the probability of the infected user S_i. P(UIN) and P(IN) are the probability of the infected and uninfected user having the symptoms of S_i. Depends on the probability the category of the user is identified, the maximum probability classifies the user[23].

Here the classification of the user is identified using the NBN algorithm,

Input: Presume, ZVD Symptoms S_i and Ref No giving n no of values as input.

Output based on the symptoms categorization is identified as IN or UIN based on the symptoms

Algorithm:

Initialization: the refno and S_i;

If (Ref no=n), then, record already exist;

Update the details of the user

Else

Create a new record and store it in DB;

If Ref no (updated data=old data)

Then,

Store the values in the existing data;

Else

Store the classification value in DB;

Update the patient's category

If S_i matches with the a_i, where i>=n(no of patients)

Then

Class C=IN;

Else

Class C=UIN;

Update category of the patient;

Alert send to doctor, user and hospitals;

End if.

The maximum probability of the infected and uninfected are examined using the NBN algorithm. After initialization of the symptoms the user details are given an updated in the system. Every user got a ref no, using that the details are updated periodically. Using the NBN classifier if any user matches with the symptoms, then he found to be infected else not. Then, using the communication system a message alerts will be send to the concerned person.

Health Communication

The health communication is a major part of the model. The communication between the user and the health sector are communicating through this part. After receiving all the required details the alert message is given to the user based on the symptoms matching whether he is infected or not. So that a precaution and preventive steps to be taken by the use. And also based on the mosquitoes breeding to avoid the growth of it preventive measures are sent frequently to the user [17] [18]. These are for reducing the risk factor of ZVD. Periodically the symptoms are checked with the user and if there is an infection found an alert message is sent to the nearby hospitals and clinics, government sectors etc [15]. So that the government found the risky place where the risk is high depends upon the data collected and analyzed for preventing from mosquitoes breeding. In the data fragmentation technique the priority is given to the sensitivity of the data with the help of TPA [20]. The health communication comes in the last level as it is common to all to know about the symptoms.

IV. EXPERIMENT AND PERFORMANCE ANALYSIS

In our proposed system, the experiment is divided as segments. Initially the synthetic data is generated. Then the training and testing of NBN are conducted using the Weka tool. The NBN classifier categorizes the infected and uninfected user with the help of synthetic data.

Synthetic Data

The symptoms of the Zika virus are analyzed completely for the proposed system to make the testing. Since it is rare in India for getting the details of the patients, we are planned to use the synthetic data where we can able to have all the combination as per our assumption. Here seven different symptoms are used for the diagnosis of infection. In the below Table 2, the possible combinations of the ZVD symptoms samples are given [20].

Table 2: Combination of ZVD symptoms using Sample data

S.No	Fever	Joint pain	Conjunctivitis	Skin rashes	Muscle pain	Headache	Exposure to risk prone area
1	Y	Y	Y	Y	N	N	N
2	Y	N	Y	N	Y	N	N
3	Y	Y	Y	Y	N	N	Y
4	Y	N	N	Y	N	N	Y
5	Y	Y	N	Y	Y	Y	N
6	Y	N	N	Y	Y	Y	Y
7	Y	N	Y	Y	N	N	Y
8	Y	Y	N	N	N	Y	Y

Then a synthetic data of mosquito breeding location and site of mosquito dense site are generated. So the mapping is done randomly with the location, symptoms along with the user. So it is easy to identify the infected and uninfected patient as well as the preventive measures to be taken by the government sectors and hospital individuals [21].

In the NBN training set the data is used as a training data set and it shows the detailed accuracy level [22] [23]. The details of category wise accuracy result taken from Weka.

Category Wise Accuracy Results in Weka

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	1.000	0.364	0.778	1.000	0.875	0.903	I
	0.636	0.000	1.000	0.636	0.778	0.903	U
Weighted Avg.	0.840	0.204	0.876	0.840	0.832	0.903	

Using the NBN classifier the accuracy is 88 percent. True positives are considered to be correctly classified instance where false positive are incorrectly classified instances. The values of the TP and FP are 0.840 and 0.204 respectively for minimal dataset. The significance of the result is found through the precision and recall parameters which give the result as 0.876 and 0.840 respectively. Next the parameter F-Measure and ROC area gives the classifier accuracy with the values of 0.832 and 0.903. From the result our proposed NBN algorithm is justified with the accuracy of 88 percent.

Table 3: Comparative analysis of Different Classification Algorithms

Author	Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)
Koc et al. 2014	LR	65	68	69
Yan et al. 2006	MLP	77	75	80
Hagan et al. 1996	NN	83	84	85
Proposed methodology	NBN	88	80	73

In Table 3 the different classification algorithms are discussed, the Linear Regression classifier [17] produced the accuracy level of 65%, Sensitivity level 68% and specificity 69%. Next, the Multilayer Perceptron algorithm [18] shows the result of accuracy 77%, sensitivity 75%, and specificity 80%. Then there comes the Neural Network classifier [19] which gives accuracy level as 83%, sensitivity 84%, and specificity 85%. Finally the proposed algorithm NBN shows a result of accuracy 88%, sensitivity 80%, and specificity 73%.

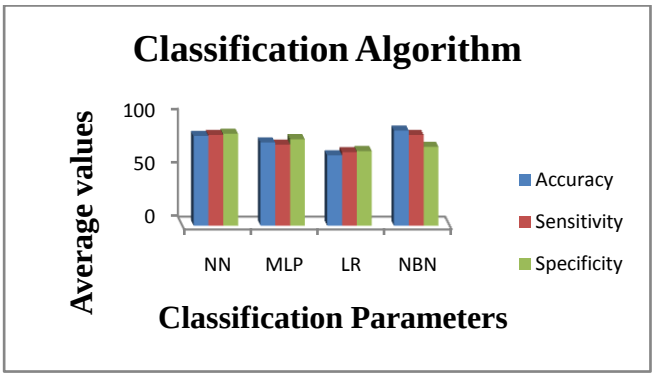

Figure 3: Performance analyses of Classification Algorithms

The Figure 3 shows the comparative graph analysis for the performance of the algorithms. The Accuracy, sensitivity, specificity of the algorithms of the existing and the proposed are explained. Comparing with the result the NBN performance level is higher than the other. The sensitivity level is higher between NN and NBN comparing with LR and MLP.

The Specificity level is higher in NN and comparatively low in LR. LR algorithm is low in every parameter and the average values of the performance and the accuracy is good at NBN.

V. CONCLUSION

ZVD is a major issue in the healthcare services, which spreads globally. Required steps should be taken for preventing and protecting public from ZVD. Here a cloud based system using mobile phones and IoT for detecting and preventing the ZVD. NBN classifier is used to prognosticate ZVD patient whether infected or uninfected with the help of their symptoms. The sensors used to obtain the details about the mosquito breeding sites, dense location everything are observed. The GPS is used to map the location of the sites specified by the sensor and finding the risk prone area. The proposed method produces the accuracy of 88 percent using the classifier. It insists the government healthcare unit to protect the area from dense mosquito. It intimate the concern user whether infected or not and also the hospitals, doctors. In future work the GPS site mapping, data security and some other algorithm to provide better result than NBN.

REFERENCES

1. Sarmiento-Ospina A, Vsquez-Serna H, Jimenez-Canizales CE, Villamil-Gmez WE, Rodriguez-Morales AJ. Zika virus associated deaths in Colombia. *Lancet Infect Dis*. 2016;16:523-524.
2. World Health Organization, Zika outbreak: WHO's global emergency response plan. 2016.
3. Musso D, Roche C, Robin E, Nhan T, Teissier A, Cao-Lormeau VM. Potentialsexual transmission of Zika virus. *Emerg Infect Dis*. 2015;21:359- 361.
4. Paixao ES, Barreto F, Teixeira GM, Costa CM, Rodrigues L. History, epidemiology, and clinical manifestations of Zika: A systematic review.*Am J Public Health*. 2016;106:606-612.
5. Nishiura H, Mizumoto K, Rock KS, Yasuda Y, Kinoshita R, Miyamatsu Y. A theoretical estimate of the risk of microcephaly during pregnancy with Zika virus infection. *Epidemics*. 2016;15:66-70.
6. Petersen E,Wilson ME, Touch S, et al. Rapid spread of Zika virus in the Americas - Implications for public health preparedness for mass gatherings at the 2016 Brazil Olympic Games. *Int J Infect Dis*. 2016;44:11-15.
7. Lopez-Barbosa N, Gamarra JD, Osma JF. The future point-of-care detection of disease and its data capture and handling. *Anal Bioanal Chem*. 2016;408:2827-2837.
8. Quwaider M, Jararweh Y. A cloud supported model for efficient community health awareness. *Pervasive Mob Comput*. 2016;28:35-50.
9. Mamun KAA, Alhussein M, Sailunaz K, Islam MS. Cloud based framework for Parkinsons disease diagnosis and monitoring system for remote healthcare applications. *Future Gener Comput Syst*. 2017;66: 36-47.
10. Zhang Z, Wang H, Wang C, Fang H. Cluster-based epidemic control through smartphone-based body area networks. *IEEE Trans Parallel Distrib Syst*. 2015;26:681-690.
11. Sareen S, Sood SK, Gupta SK. Towards the design of a secure data outsourcing using fragmentation and secret sharing scheme. *Information Security Journal: A Global Perspective*. 2016;25:39-53.
12. Mahalakshmi, B., and G. Suseendran. "Effectuation of Secure Authorized Deduplication in Hybrid Cloud." *Indian Journal of Science and Technology* 9.25 (2016).
13. Hadavi, A. M., & Jalili, R. (2010). Secure data outsourcing based on threshold secret sharing; towards a more practical solution. In Proceedings of VLDB PhD workshop, Singapore.
14. Han, J., Susilo, W., & Mu, W. (2013). Identity-based data storage in cloud computing. Future Generation Computer Systems, 29(3), 673–681. doi:10.1016/j.future.2012.07.010
15. Sareen S, Sood SK, Gupta SK. A cloud-based seizure alert system for epileptic patients that uses higher-order statistics. *IEEE Comp Sci Eng*. 2016;18:56-67.
16. Chen, D., & Zhao, H. (2012). Data security and privacy protection issues in cloud computing. In Proceedings of the ICCSEE 2012, pp. 647–651, Hongzhou.
17. Ko, M., and A. Barkana. 2014. "Application of Linear Regression Classification to Low-Dimensional Datasets." Neurocomputing 131: 331–335. doi:10.1016/j.neucom.2013.10.009.
18. Yan, H., Y. Jiang, J. Zheng, C. Peng, and Q. Li. 2006. "A Multilayer Perceptron-Based Medical Decision Support System for Heart Disease Diagnosis." Expert Systems with Applications 30 (2): 272–281. doi:10.1016/j.eswa.2005.07.022.
19. Beale, Mark, Martin T. Hagan, and Howard B. Demuth. "Neural network toolbox." *Neural Network Toolbox, The Math Works* 5 (1992): 25.
20. Mahalakshmi, B., and G. Suseendran. "An Analysis of Cloud Computing Issues on Data Integrity, Privacy and Its Current Solutions." *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 839, 2019, pp.467-482, Doi: 10.1007/978-981-13-1274-8_35.
21. Sareen, Sanjay, Sandeep K. Sood, and Sunil Kumar Gupta. "Secure internet of things-based cloud framework to control zika virus outbreak." *International journal of technology assessment in health care* 33.1 (2017): 11-18.
22. Sareen, Sanjay, Sunil Kumar Gupta, and Sandeep K. Sood. "An intelligent and secure system for predicting and preventing Zika virus outbreak using Fog computing." *Enterprise Information Systems* 11.9 (2017): 1436-1456.
23. COE, JCDM. "Performance comparison of Naïve Bayes and J48 classification algorithms." *International Journal of Applied Engineering Research* 7.11 (2012): 2012.