Application Of Self-Organizing Maps For Prioritization Of Malaria Control Operations In Changlang District, Arunachal Pradesh

U Murty, N Arora

Citation

U Murty, N Arora. *Application Of Self-Organizing Maps For Prioritization Of Malaria Control Operations In Changlang District, Arunachal Pradesh*. The Internet Journal of Epidemiology. 2006 Volume 4 Number 2.

Abstract

Arunachal Pradesh is highly endemic for malaria. Despite of all the efforts under NMEP, malaria continues to pose a serious threat. Hence, there is a need for an alternate technology for efficient targeting of malaria in this region. Here, we present application of Self-Organizing maps (SOM), a powerful data mining technique for prioritization of control operations.

INTRODUCTION

Malaria is the second most fatal communicable public health problem in 90 countries in the world (1) and causes more than 300 million acute illnesses and at least one million deaths annually (2). The annual malaria burden in India estimates nearly 2 to 2.5 million cases. North-Eastern region of India is in the Indo-Chinese hill zone of Macdonald's classification of stable malaria (3) and constituting 09% of total malaria cases in India (4). In North-Eastern States efficient malaria transmission is maintained during most months of the year.

Even though several anti-malaria programmes are being implemented under National Malaria Eradication Programme newly termed as National Vector Borne Diseases Control Programme since 1958 but this region continues to face the malaria threat (5, 6). The major problems are predominance of Plasmodium falciparum, chloroquine resistance (7), highly efficient multiple resistant vectors (8), congenial eco-climatic conditions (9, 10), lack of skilled manpower and ineffective communication between health researchers and policy makers. Apart from the above reasons, difficult terrain, frequent floods and improper road communication are unavoidable constraints for the implementation of anti-malarial operations in Arunachal Pradesh (11). Hence there is an urgent need for exploitation of data mining tools, which can prioritize the malaria endemic areas.

Data mining is a technique of Artificial Intelligence (AI), which digs out the hidden knowledge from the massive heaps of data. Cluster Analysis has been widely applied in various fields like Breast and colo–rectal cancer (12), diabetes mellitus for analyzing immune complexities (13) clinical chemical diagnosis of liver diseases (14), neurological diseases (15) etc. In accordance with the important role of Data mining applications, Self Organizing Maps (SOM) were used in the present study to prioritize the malaria endemic zones for disease management.

MATERIALS AND METHODS

STUDY AREA

Changlang district of Arunachal Pradesh was selected for the study, which has been identified as malaria endemic zone due to hilly terrain, climatic conditions, dense tropical forests, which make it an ideal abode for vector (9).

Changlang District lies in the South-Eastern corner of Arunachal Pradesh stretched to an area of 4662 sq. Km. The District lies between the latitudes 26°40'N and 27°40'N, and longitudes 95°11'E and 97°11'E. It is bounded by Tinsukia District of Assam and Lohit District of Arunachal Pradesh in the North, by Tirap District in the West and by Myanmar in the South-East. The major rivers are Noa-Dehing, Namchik and Tirap. Climatic conditions and rainfall are very much influenced by the terrain. The annual rainfall ranges from 3800 mm to 4866 mm. The major rainfall is received during June to October. Tropical rain forest forms major bulk of vegetation. The aboriginal inhabitants of Changlang district are the Tangsas, Singphos and Tutsas. Platform (Machang) types of houses are generally preferred in all tribes and subtribes having more or less same pattern. The main occupations of the indigenous people of the district are agriculture and allied activities and animal husbandry.

RAW DATA COLLECTION

Raw data was collected from the Directorate of Health, Govt. of Arunachal Pradesh, which consists of epidemiological and entomological aspects of malaria.

PARAMETERS USED

EPIDEMIOLOGICAL PARAMETERS

Parameters like Monthly Parasite Incidence (MPI), Slide Positivity Rate (SPR), and Slide falciparum rate (SFR) were considered.

ENTOMOLOGICAL PARAMETERS

Parameters such as Indoor human dwelling collection, outdoor human dwelling collection, whole night bait collection on indoor human bait, whole night bait collection on outdoor human bait, whole night bait collection on indoor animal bait, indoor cattle resting collection, Indoor cattle biting collection, larval Collection using per dip capture method and parity rate of 4 vector species were used for this study.

DATA MINING –SELF –ORGANIZING MAPS

Artificial Intelligence is a branch of computer science that studies how to endow computers with capabilities of human intelligence. The term was coined in 1956 by John McCarthy at the Massachusetts Institute of Technology. Self-organizing map (SOM) was invented by Professor Teuvo Kohonen, Helsinki University of Technology, Finland during 1960s (Self Organizing Maps).

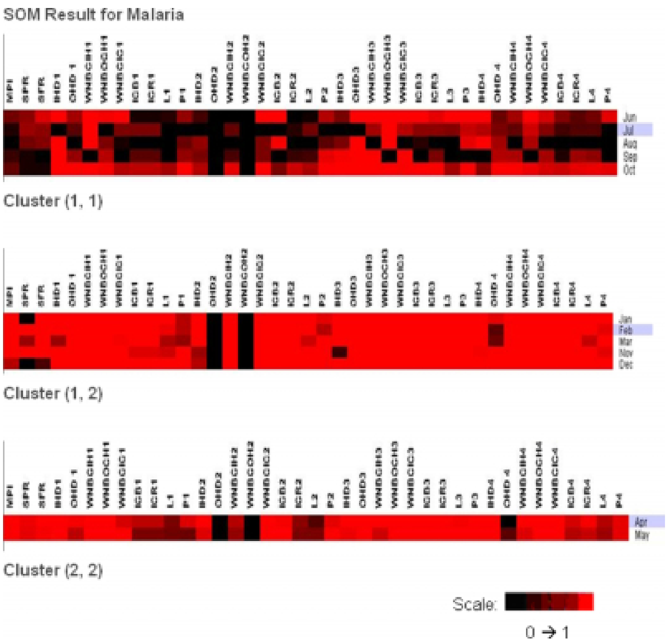
In SOM the neurons are organized in a lattice, usually a one or two-dimensional array, which is placed in the input space and is spanned over the inputs distribution. Using a two-dimensional SOM network, it is possible to obtain a map of input space where closeness between units or clusters in the map represents closeness of the input data. Processing units in the SOM lattice is associated with weights of the same dimension of the input data. Using the weights of each processing unit as a set of coordinates the lattice can be positioned in the input space. During the learning stage the weights of the units change their position and “move” towards the input points. This “movement” becomes slower and at the end of the learning stage the network is “frozen” in the input space. After the learning stage the inputs can be associated to the nearest network unit. When the map is visualized the inputs can be associated to each cell on the map. One or more cell that clearly contains similar objects can be considered as a cluster on the map. These clusters are generated during the learning phase without any other information. It is not necessary to supply to the network cluster prototypes or examples. The main applications of the SOM is the visualization of high-dimensional data in a two dimensional manner, and the creation of abstractions like in many clustering techniques. The characteristic that distinguishes the SOM net from the other classification algorithms is that not only similar inputs are associated to the same cell but also neighborhood cells contain similar documents. This property together with the easy visualization makes the SOM map a useful tool for visualization and clustering of large data sets.

RESULTS

Normalized data on clustered using SOM yielded 3 clusters on a 2x2 grid shown in figure 1. Unsupervised learning was done on the fly using the data using a learning constant of 0.01 and for 10,000 iterations following which the data got clustered with the first cluster (1, 1) showing relatively minimum risk during the months of June, July, August, September, and October and as we proceed towards the last cluster (2, 2) we find regions with high risks. These Self Organizing Maps can help in cost effective vector control by spatial and temporal targeting.

Figure 1

Figure 1: showing clusters obtained on 2X2 grid where MPI=Monthly Parasite Incidence, SPR=Slide Positivity Rate, SFR=Slide Falciparum Rate, IHD=Indoor Human dwelling collection. OHD=Outdoor human dwelling Collection, WNBCIH=Whole Night Bait Collection on Indoor Human, WNBCOCH= Whole Night Bait Collection on Outdoor Human, ICB=Indoor Cattle Biting collection, ICR=Indoor Cattle resting collection and 1 = , 2 =, 3 = , 4 =



DISCUSSION

Cluster (1, 1): This cluster has moderate MPI as well as SPR and a decline in vector abundance as reflected in different mosquito collections. Larval abundance was also observed to be at a low level in monsoon season due to heavy precipitation received during these months. This region requires both drug administration and vector control measures at relatively less priority.

Cluster (1, 2): It has been observed that in cluster , very high Malaria Parasite Incidence and Slide Positivity Rate were observed from January to March(pre-monsoon season) and November and December (post monsoon season).Anopheles phillipiensis was clearly found to have endophilic and endophagic tendencies while other major vectors found in significant shows ambivalence in their resting and feeding behavior. Throughout the pre monsoon and post monsoon season, a very high landing rate, an indicative of man vector contact, crucial for malarial transmission was observed, resulting in a high degree of malaria transmission. Larval index was observed to be very high in this cluster, showing a slight decline only in the month of March. Vector control seems to be an efficient weapon to reduce transmission with direct consequences in lowering the incidence of malaria cases. Hence, the months that are clustered in this group are high priority months for control operations for decreasing man vector contact by focusing on anti larval measures specifically as well as indoor and outdoor operations keeping in mind ambivalent resting and feeding habits of major malarial vectors in this region.

Cluster (2, 2): The months included in this cluster showed a very high level of malaria transmission as indicated by malariometric parameters. The larval abundance showed a clear decline because of flushing away of breeding grounds on onset of rain. We propose that the months in cluster (1, 2) and (2, 2) may be the most favorable seasons in which to envisage efficient and efficacious anti-malarial activities. This region requires both drug administration and vector control measures at relatively high priority with relatively lower emphasis on anti larval measures.

CONCLUSIONS

Changlang district is highly endemic for malaria. Effective control of malaria in this region of perennial transmission requires effective use of information technology tools to bridge the communication gaps between health workers and policy makers. Here we show the use of self organizing maps as valuable tool in predicting months and mode of action specific for efficient vector control operations in the region.

ACKNOWLEDGEMENTS

Authors are grateful to Director, IICT, Hyderabad for his continuous support and encouragement. Neelima Arora thanks CSIR for Senior research Fellowship. Thanks are due to Department of Science and Technology, Government of India for funding the project.

References

1. World Health Organization (WHO), Expert Committee on Malaria.WHO Expert Committee on Malaria, Twentieth Report. Geneva, WHO (1998).
2. World Health Organization (WHO). The World Health Report, 1999: WHO, Geneva. (1999).
3. MacDonald G, The epidemiology and control of malaria. Oxford University Press, London (1957).
4. Shiv Lal et al., Journal, Indian Academy of Clinical. 5(1):19 -23(2000).
5. PK Mohapatra et al., ICMR Bulletin 28 No 3 21-30(1998).
6. PK Sen. Indian J Public Health, 38(4):155-8.(1994)
7. PK Mohapatra et al ., Indian Journal of Medical Research, (2003)
8. Sharma, V. P. Ind. J. Med. Res. 103 : 26-45(1996)

9. V Dev et al. , JAPI 51(2003)
10. R L Yadava & RS Sharma, J Commun Dis.27 (4):262-6(1995)
11. PK Mohapatra et al., J. Natl Med J India.14(3):139-42.(2001)
12. P.K. Newby & Katherine L. Tucker Nutrition Reviews, ,62(5):177-203 (2004)
13. Francisco J. Quintana et al., Journal of Autoimmunity 21(1) :65-75 (2003)
14. Folkerts et al. , J Clin Chem Clin Biochem. 28(6):399-406 (1990).
15. Davide Seripa et al., Neurobiology of Aging, 26(4):455-464 (2005).
16. Teuvo Kohonen, Self-Organizing Maps, Springer Series in Information Sciences, (30),Springer, Berlin, Heidelberg, New York, 1995, 1997, 2001, 3rd edition.

Author Information

U. S. N. Murty, Ph.D.
Deputy Director, Scientist F, Biology division, Indian institute of Chemical Technology

Neelima Arora, M.Sc.
Senior Research Fellow, Bioinformatics Group, Biology Division, Indian Institute of Chemical Technology