IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

Efficiency of Classification Techniques in Detecting Congestion in Wireless Sensor Networks

Mrs. Jayashri B. Madalgi [1] and Dr. S. Anupama Kumar[2]

Research scholar, VTU, Gogte Institute of Technology, MCA Department, Belgaum, India[1]

RV College of Engineering, Department of MCA, Bangalore, India[2]

Abstract: Wireless Sensor Network (WSN) is network of hundreds/thousands of sensor nodes. Each node is capable to sense, process and transmit the environmental information. Congestion occurs in wireless sensor networks when an event occurs. Congestion leads to performance degradation of a system. The data mining techniques help to detect congestion and then it can be mitigated by adjusting transmission rate. In this paper we analyze the efficiency of data mining classification to detect the congestion in WSN. In this paper we have implemented PART, RIPPER and j48 classification algorithms to detect the congestion over the network. For the given data set, it is found that PART algorithm is more accurate than RIPPER and j48 in detecting the congestion.

Index Terms: WSN, Data mining, Congestion control, classification, PART, RIPPER, j48 algorithm.

I.INTRODUCTION

The advances in wireless communication and microelectronic devices have developed the low-power sensors and ease in deployment of large-scale sensor networks.[9] Sensor networks has many applications in most of the domains such as habitat monitoring, object tracking, environment monitoring, military, disaster management. With the growth of networks and increased link speeds .congestion in networks is a significant. Congestion occurs when the packet traffic load in the network is greater than the capacity of the network. It is a situation in which too many packets are present in a part of the subnet, thus leading to performance degradation. [2]

1.1 Congestion In Sensor Networks

The traffic in WSN can be divided into 2 streams. The downstream is from the sink/base station to the sensors and the upstream is from the sensors to the sink/base station [12]. The downstream traffic is of one-to-many where as upstream is many-to-one nature[2]

1.2: Nature Of Congestion

In WSN applications congestion can be differentiated into 2 categories: the congestion that arises near the sink and congestion that arises near the sources.[14]

Congestion near the event-sources:
Occurrence of an event results in traffic burst from sensor nodes near the event area, leading to collisions and loss in packets at the sources.

Congestion near the sink:
The traffic generated at multiple source nodes travels in multi hop fashion towards the base station, The data is generated when an event is detected at the same a time by the source nodes .It results in more packet flow in the region near base station due to the funnel-like communication pattern.

This increase in packet flow as traffic results in congestion .In section 2 we analyze the drawbacks of network protocols. In section 3 the results of implementation of classification techniques are shown and in section 4 the results are analyzed.

II. RELATED WORK

Many techniques exists which are specifically invented for the wireless sensor networks. These protocols are deployed by the layers of the WSN OSI stack. [2]

A. Techniques used by Data Link Layer

In wireless sensor networks mainly there are two congestion types - channel collision and buffer congestion. This layer overcome channel collision by using the following mechanisms CSMA-Carrier Sense Multiple Access,FDMA-Frequency Division Multiple Access,TDMA-Time Division Multiple Access.

B. Techniques used by Network Layer Techniques

BOBRED- Beacon Order Based Random Early Detection-active queue management techniques are effective in a limited network with few sensors and intermediate devices (routers). [9]

C. Transport Layer Techniques

The Protocol developed by the IETF (Datagram Congestion Control Protocol -DCCP) the standard was accepted in year 2006. [12].

D. Techniques with Cross Layer Nature

This technique combines mechanisms of different layers of the network. Hop-by-hop flow control used by transport layer, source traffic with limited rate and prioritized data link layer[2].The network protocols mechanisms provide congestion control by concentrating on buffer length and channel capacity. These protocols do not provide packet recovery mechanism. This is essential for critical applications where minimal packet loss is required. There is possibility of loss of control information packets in network because of buffer overflow and busy channel. All the research work on protocols is carried out using simulators .Implementing these in real application with huge sensor networks is difficult. The data mining techniques do not generate control information.

III. IMPLEMENTATION OF CLASSIFICATION TECHNIQUES

Machine learning techniques are also applied to congestion control over wired and wireless networks. The authors of [4],[5] and [7] have analyzed the status of wired and wireless network using the parameters like routing algorithm used ,traffic load, queue length, bandwidth, route available and predict possibility of congestion and provide alternate route for congestion control[5].

The classification algorithms are applied over WSN to detect/predict congestion. Decision tree is the most popular supervised learning algorithms. The pruned trees are small at each intermediate step and the computation steps of a prediction are fast which consists a sequence of comparisons along the path of tree.The database used is from the network traffic data in wireless sensor network application.

The parameters used are queue length, packet loss, packet service time, packets inter arrival time and delay. The delay of a packet is the delay between the time packet was sent and the received time . The packet inter arrival time is the difference between the time packet is sent and the time the previous packet was sent .These parameters value are dependent on the state of congestion in network and thus it makes it possible the occurrence of congestion to be predicted by using these parameters for a number of packets. The data set consist of ARFF(attribute relation file format) data file with attributes queue length, ,delay, packet interartime, packet service time, packet loss and classification attribute class {con,nocon}con denotes congestion and nocon denotes no congestion.

3.1 PART Algorithm

Eibe Frank and Ian H. Witten have developed a PART algorithm .It builds partial tree in each iteration and construct rules from best leaf of decision tree. It adopts separate and conquer strategy. A decision tree is built using current set of instances and the leaf with most coverage is used to build rule and then tree is discarded. [22]

PART decision list:
queuelength<= 15 AND packetloss <= 2
AND queuelength > 10: nocon (23.0/2.0)
packetinterartime <= 4: con (38.0/3.0)
delay > 1 AND queuelength <= 11: nocon (7.0): con (5.0)
Number of Rules : 4

The confusion matrix result is as shown below:

TABLE I: CONFUSION MATRIX OF PART ALGORITHM

con	nocon	class
40(True)	2 (False)	con
3(False)	28 (True)	nocon

3.2 RIPPER Algorithm

William H .Cohen developed RIPPER algorithm.In this the rules induction is using a sequential covering algorithm .we have applied this algorithm for the data set .Rules are learned for one class at a time .[23]
The results are shown below.

JRIP rules:
(packetloss <= 1) and (queuelength <= 13)
=> class=nocon (17.0/1.0)

(queuelength <= 12) and (packetloss >= 7)
=> class=nocon (3.0/0.0)

(packetloss <= 2) and (delay <= 2) and
(queuelength >= 11) => class=nocon (7.0/0.0)

=> class=con (46.0/5.0)
Number of Rules : 4

The Confusion Matrix result is as shown below:

TABLE II:CONFUSION MATRIX OF RIPPER ALGORITHM

con	nocon	class
41(True)	1(False)	con
5 (False)	26(True)	nocon

3.3 J48 Algorithm

J48 algorithm is an open source and it is s java implementation of C4.5 algorithm[21] in Weka data mining tool. In 1980s, J Ross Quinlan, a researcher developed c4.5 a decision tree algorithm which is a successor of ID3. A decision tree is built in top down manner using recursive divide **-** conquer approach.[24]

J48 pruned tree
queuelength <= 15
| packetloss <= 2
| | queuelength <= 10
| | | delay <= 1: con (3.0)
| | | delay > 1: nocon (3.0)
| | queuelength > 10: nocon (23.0/2.0)
| packetloss > 2
| | packetinterartime <= 4: con (20.0/3.0)
| | packetinterartime > 4: nocon (5.0/1.0)
queuelength > 15: con (19.0)
Number of Leaves : 6

The confusion matrix result as shown below:

TABLE III CONFUSION MATRIX OF J48 ALGORITHM

con	nocon	class
39(True)	3(False)	con
3(False)	28(True)	nocon

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

The decision tree generated to classify congestion is as shown below:

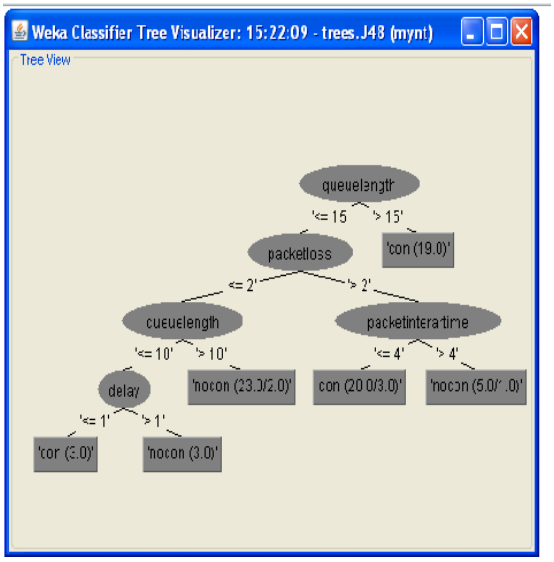

Fig 1. Decision tree of j48 algorithm

After analyzing the results it can be stated that queue length and packet loss are 2 important attributes and the value less than threshold indicates no congestion in network.

IV. CONCLUSION

The different data mining classification algorithms PART, Ripper and j48 algorithms can be applied to detect congestion in wireless sensor networks. The incorrectly classified instances are 5 in PART and in RIPPER and J48 algorithm. are 6.

TABLE IV: COMPARISON OF 3 ALGORITHMS

Name of the Algorithm	Accuracy in Detecting Congestion (%)	Time taken (seconds)
PART	93.15	0.02
RIPPER	91.78	0
j48	91.78	0

The time taken to build model is almost same in all algorithm. However Based on confusion matrix and it can be clearly seen that accuracy of PART (93.1507%) is highest. Therefore it can be suggested that PART is most appropriate classifier to detect congestion in WSN. PART classifiers can be used as predictive model for congestion detection in wireless sensor network more effectively.

REFERENCES

[1]. Ekaterina Dawhkova and Andrei Gurtov "Survey on congestion control mechanisms for Wireless Sensor networks "center for wireless communication,finland,2013
[2]. C. Wang.B. Li,K Sohrarby and M. Daneshmand " UPstream Congestion control on wireless sensor networks through cross layer optimization" IEE journal on selected areas in communications, VOL 25,No. 4,May 2007.
[3]. L. Breiman, J. Friedman, R. Olsen, and C. Stone. "Classification and Regression Trees." Wadsworth International (California)
[4]. Pierre Geurts, Ibtissam El Khayat, Guy Leduc," A Machine Learning Approach to Improve Congestion Control over Wireless Computer Networks"
[5]. Sasan Adibi ,"Data Mining - A Captured Wired Traffic Approach", International Journal of Advanced Science and Technology Vol. 21, No. 2, Aug, 2010
[6]. Soukaena Hassan Hasheem"An Association Rules Analysis to enhance solving the Congestion Problem", Journal of Emerging Trends in Computing and Information Sciences Volume 2 No.7, JULY 2011
[7]. Manoj Devare and Ajay Kumar , "Clustering and Classification (Time Series analysis) Based Congestion Control algorithm: Data Mining Approach", IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.9, September 2007 pp 241-246
[8]. Sudip Misra, Isaac Woungang, Subhas Chandra Misra ,"Guide to Wireless Sensor Networks"
[9]. Ameer Ahmed Abbasi, Mohamed Younis "A survey on clustering algorithms for wireless sensor networks"june2007." Springer publications
[10]. Shivangi Borasia and Vijay Raisinghani, "A Review of Congestion Control Mechanisms for Wireless Sensor Networks "
[11]. Md. Abdur Rahman, Abdulmotaleb El Saddik and Wail Gueaieb, " Wireless Sensor Network Transport Layer:State of the Art " Springer-Verlag Berlin Heidelberg 2008
[12]. Md. Mamun-Or-Rashid, Muhammad Mahbub Alam, Md. Abdur Razzaque,and Choong Seon Hong "Reliable Event Detection and Congestion Avoidance in Wireless Sensor Networks"© Springer-Verlag Berlin Heidelberg 2007
[13]. Vikram P. Munishwar, Sameer S. Tilak, and Nael B. Abu-Ghazaleh" Congestion and Flow Control in Wireless Sensor Networks "© Springer-Verlag London Limited 2009
[14]. Tan, Steinbach, Vipin kumar "Introduction to Data mining"
[15]. Han J., and Kamber, M. "Data Mining: Concepts and Techniques". Morgan Kaufmann, 2006.
[16]. Mohammad Abu Alsheikh, Shaowei Lin, Dusit Niyato and Hwee-Pink Tan"Machine Learning in Wireless Sensor Networks: Algorithms, Strategies, and Applications
[17]. Qi Li, Zongwu Ke, Duanfeng Xia, Yuxia Sun "A Routing Protocol for Wireless Sensor Networks with Congestion Control"
[18]. Hemmat Sheikhi Mousa Dashti Mehdi Dehghan "Congestion Detection for Video Traffic in Wireless Sensor Networks" 2011 IEEE
[19]. Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten (2009); The WEKA Data Mining Software: An Update; SIGKDD Exlorations, Volume 11, Issue 1.
[20]. Prabu Rajkumar P, Arul Treesa Mathew, Sruthi N Paul, Sujitha B Cherkottu "Congestion Control in Healthcare Wireless Sensor Networks- A Data Centric Approach"Technology Research (IJSETR) Volume 2, Issue 7, July 2013
[21]. https://en.www.wikipedia.org/wiki/c4.5 algorithm
[22]. Eibe Frank,Ian H. witten "Generating Accurate rule sets without global optimization" Fifteenth International Conference on Machine Learning, 144-151, 1998
[23]. William H. Cohen "Fast Effective Rule Induction" Proceedings of the Twelfth International Conference on Machine Learning,1995
[24]. C4.5: by J. Ross Quinlan.. Programs for M Publishers, Inc., 1993achine Learning Morgan Kaufman