International Journal of Computer Applications (0975 – 8887)
Volume 30– No.9, September 2011

Efficient Flooding for a Large Sensor Networks using Network Coding

Nishant Jain
School of information technology, RGPV
Bhopal, India

Sanjeev Sharma
School of information technology, RGPV
Bhopal, India

Santosh Sahu
School of information technology, RGPV
Bhopal, India

ABSTRACT
Sensor network facilitates monitoring and controlling of physical environments. These wireless networks consist of dense collection of sensors capable of collection and dissemination of data. They have application in variety of fields such as military purposes, environment monitoring etc. Typical deployment of sensor network assumes central processing station or a gateway to which all other nodes route their data using flooding. This causes congestion at central station and thus reduces the efficiency of the network. In this work we will propose a better flooding technique using network coding to reduce total number of transmission in sensor networks resulting in better efficiency.

Keywords
Network Coding, Wireless Sensor Network, Wireless Flooding.

1. INTRODUCTION
Wireless sensor networks (WSNs) are one of the most popular networks used in computer science. It consists of some autonomous sensors used to monitor various natural activities. These natural activities may be pressure, temperature, sound vibration etc.. Sensor nodes collect data from natural activities and pass it to central processing station or gateway called sink [1]. Lots of researches have been made in various areas to enhance the overall throughput of WSNs. Development of WSNs actuated by some important applications like military surveillance's, medical sciences, natural disaster control, etc. The topology of WSNs consists of multiple sources and single sink. This causes some common problems like congestion at sink, limited resources, etc. Because of limited resources flooding is also a major problem for WSNs. Efficient flooding helps to increase the overall throughput of the WSNs.

1.1 Wireless Flooding
Flood is a communication primitive that can be used by the base station of a sensor network to send a copy of a data message to every sensor in the network. The execution of a flood starts by the base station sending a copy of the data message to every one of its neighboring sensors [2]. Whenever a sensor receives a data message, it keeps a copy of the message and forwards the message to every one of its neighboring sensors and the cycle repeats. In WSNs, flooding is a protocol that broadcast messages to other nodes [12]. Flooding is fundamental operation for time synchronization, group formation, node localization and routing tree formation. Existing flooding algorithm have demonstrated their effectiveness in achieving communication efficiency and reliability in WSNs [13]. Enhanced performance can be achieved by using network coding.

1.2 Network Coding
Network coding is the technique which is extensively used in wired networks, ad-hoc networks, and distributed sensor networks, etc. Network coding is quite different from traditional communication. Network coding achieves vast performance gains by permitting intermediate nodes to carry out algebraic operations on the incoming data [10]. Network coding allows the packets to encode and further forward it. The destination sink decodes the packets. Encoding is simply XOR of data packets which will be called as encoded packet. Decoding is XOR of data packets (except the missing one) and the encoded packet as a result the missing packet gets identified [2], [9], [10].

Consider a sensor network in Fig 1a having five nodes. Node S1, S2, S3 and S4 has some packet data to share with each and every node. Assume all links have a time unit capacity. In current approach, each node broadcasts their data and was listened by their neighboring node according to Figure. Now there is a bottleneck on node N and have 4 data packets for transmission. Node N broadcasts all 4 data packets one by one. Each node listen these broadcasts and collects their data packet. This approach requires 8 broadcasts in all.

Now consider network coding approach illustrated with the help of same example. Each node broadcasts their data which is collected by neighboring node. Now each sensor node S1, S2, S3 and S4 has 3 data packets received by their neighboring node as shown in Fig 1a. In Fig 1b and Fig 1c Node N has 4 data packets to transmit. Using network encoding approach node N encodes the data packets and broadcast it. Now all sensor nodes S1, S2, S3 and S4 listens this encoded packet and decode these packets by using network decoding shown in Fig 1d. Now this approach requires 5 broadcasts which are 37 % less than previously discussed approach. This also reduces the bottleneck, congestion at sink and total transmission on the network and in the process provides gain in bandwidth, efficiency and power resources of the nodes [14].

1.3 Deployment strategy for WSNs
Efficient deployment strategy is necessary to detect event occur in WSNs and obtain the real time data. For example for a large dense forest there no need deploy WSNs in mountain region. This can be done by deploying sub sensor networks in a distributed manner. Density of sensors depends on the occurrence of events. The positions of sensors are predetermined and position of sensor nodes identified by GPS systems. Each transmission contains a source ID and Sink ID and transmission is directed to sink node [4], [7]. Proposed topology can be viewed as subsequent part of large sensor network where each

XOR is simply exclusive-or of the packets can easily be obtained by XOR truth table. Suppose node Px and Py are two packets. Such that Px=10110 and Py=01101.Packet encoding=Px XOR Py=10110 XOR 01101= 11011=Pz. Packet decoding=Px XOR Pz=10110 XOR 11011=01101=Py and Py XOR Pz=01101 XOR 11011=10110=Px. Where Pz is encoded packet.

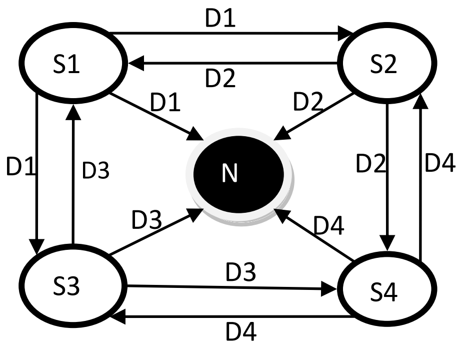


Fig 1a: Message Broadcasting

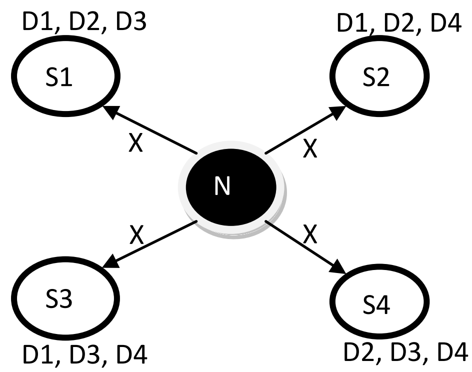
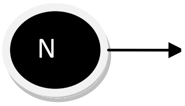


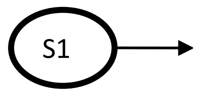
Fig 1c: Encoded Message Broadcasting



$(((D1 \oplus D2) \oplus D3) \oplus D4) = X$

$\oplus$ = XOR Function

Fig 1b: Message Encoding



$(((D1 \oplus D2) \oplus D3) \oplus X) = D4$

Fig 1d: Message Decoding

node taking part in data transmission using current communication approaches.

2 RELATED WORK

In this section we explore the history of network coding and wireless flooding. Ahlswede et al. [1] showed that with network coding, as symbol size approaches infinity, a source can multicast information at a rate approaching the smallest minimum cut between the source and any receiver. Practical deployment of network coding is thoroughly described in [5], [7], [8].Various network coding techniques like linear and random network coding are classified in [9]. [6] Provides a systematic method to quantify the benefits of using network coding in the presence of multiple concurrent unicast sessions. A robust network coding aware data aggregation approach which will result in better performance of the network by reducing the number of transmitted messages in the network is discussed in [4]. It also gives protection from link failure to many-to-many network flows from multiple sensor nodes to sink nodes. Benefits of network coding over wireless networks are described in [14]. Various efficient flooding techniques are explained in [3], [12], [13]. Also large WSNs topologies are described in [4]. This gives us motivation to implement better flooding for large WSNs using network coding.

3 PROPOSED SCHEME

In this section we will discuss system model, proposed approach and algorithm for proposed scheme.

3.1 System Model

In this work we have considered sensor nodes are deployed in a systematic manner to achieve efficient network coding. These node senses data from the environment and transmits to sink node. However these nodes are further classified into three categories. Sensor nodes, who senses data and transmit, some relay nodes who simply forward the data packet to sink node and aggregate node identifies the opportunity of network coding and performs it. The function f is used by the aggregate node to decide whether to apply network coding or not [4]. This topology is recursive. Consider Fig 2 Node 1, 2,3,4,7 and 10 are initial nodes. These nodes are simply a sink and relay nodes and it can be called as initial level L_0 and it is contains 6 nodes. Now topology is constructed recursively with Level L_1, Level L_2, Level L_3......Level L_p. Level L_1 has total 12 nodes. 4 of them are aggregate nodes and 8 of them are sensor nodes. Hence we can say Level L_1 contains (r+q) nodes where r is number of sensor nodes and q is number of aggregate nodes. Now Level L_2 can be obtained by 2(r+q) nodes. This recursive method continues to obtain a large sensor network and makes our topology scalable [4].

To get better performance and efficient network coding we have to made sum assumptions.

- Node deployment is 2D.
- Each sensor node has unique id and sink node maintains the id these sensor nodes.
- All transmission contains Source-ID and Sink-ID.
- Nodes must approximately at equal distance.
- Let p and q are the two received data by an aggregate node and $\boldsymbol{f}$ is a binomial function which computes significant difference between the two data values and returns yes or no. If p and q are differ not more than γ then the value of the function f is false else it returns

true. The absolute value of the difference is denoted by $d = p \sim q$. $\boldsymbol{f}$ $:\{ 0 \text{ if } d < \gamma, 1 \text{ else}\}$

3.2 Proposed Work

This research work would primarily concentrate on the design and development of sensor networks and its overall throughput. Focus of the application on the better flooding technique using network coding to reduce total number of transmission in sensor networks resulting in better efficiency. Specified sensor network topology is used as shown in Fig 2 to achieve better results in wireless flooding networks.

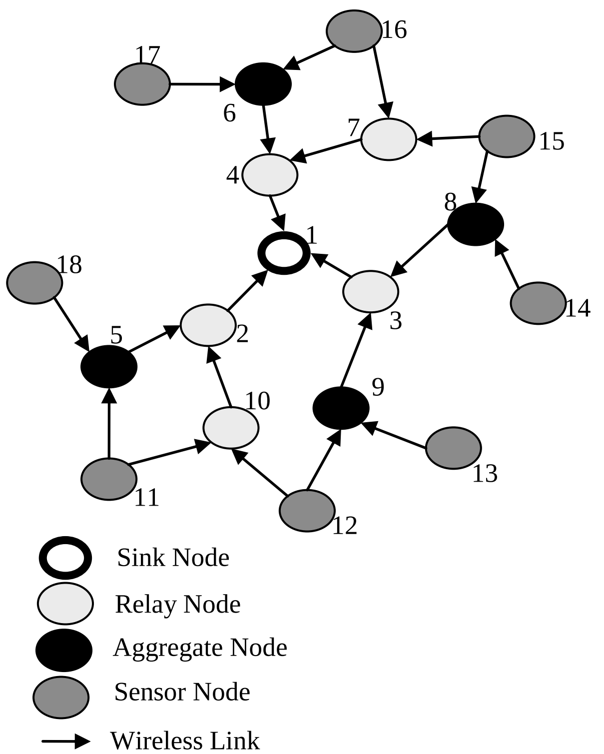


Fig 2: Sensor Network Topology

Node placement is in two dimensional as shown in Fig 2. Now node {1} is sink node. Nodes {5, 6, 8, and 9} are aggregate nodes. Nodes {2, 3, 4, 7, and 10} are relay nodes. Nodes {11, 12, 13, 14, 15, 16, 17 and 18} are sensor nodes.

For Example suppose sensor node 12 and 13 have some information to flood. Now node 12 forwards (Broadcasts) it's information to relay node 10 and aggregate node 9 and sensor node 13 forwards its information to aggregate node 9. Relay node simply forwards its information to sink node. Aggregate node 9 has two packets to forward. It encodes the packets using (XOR) technique and further forwards it to sink. Now sink has two packets. One is data packet transmitted by node 12 via relay node 10 and encoded packet transmitted by aggregate node 9. It simply decode the packet by again XORing both the packets and collect the data packet transmitted by sensor node 13. Here we can clearly see that sink node achieve the packets in 6 broadcasts instead of 8 which is currently done by wireless flooding.

Before hopping to algorithm 1 would like to put emphasis upon the point that we are not providing the answer to the question: "When to encode data?". For the purpose of this work we have used random variable to decide upon when to encode, with equal probabilities.

3.3 Algorithms

To enable encoding and decoding of packets we have used two types of packets namely normal_packet and code_packet. Where size of normal_packet is fixed and size of code_packet is will be twice of size of normal_packet plus size of header. Let at an aggregator node we need to encode pkt1 and pkt2. This is done as following: pkt1 is XORed with pkt2 and is encapsulated under new header and is then forwarded depending upon the new header. We have also included a bit in each packet namely codeOn bit which is set if packet is code_packet and unset if it is normal_packet (though it is redundant as type of packet can be identified by size) in our case. Decoding is performed by first removing the additional header and then again XORing the packet with other appropriate packets.

When aggregate node receives data from the sensor nodes function $\boldsymbol{f}$ finds the difference between the data. If difference is less than some predefined γ and is not significant then there is no need of encoding and one of the data selected at random, is forwarded to sink node. If the difference is greater than γ then aggregate node encode (XORs) data and forward to sink node. Each aggregate node caches the previously received data from the sensor nodes. If data difference between previously received data and currently received data is significant then encoding done and data transmitted to sink node. Otherwise no encoding done and the data is transmitted normally.

Algorithm: **AggregateHeuristic**(packet pkt1, packet pkt2)

//pkt_{1i}& pkt_{2i} is i_{th} packet sent by leaf node 1 and leaf node 2 respectively.

{

If $\boldsymbol{f}$(data(pkt_{1i}), data(pkt_{2i})) == 0

{

If(data(pkt_{1i}),data(pkt_{1i-1}) !=0)

{

Perform network coding on pkt_{1i} and pkt_{1i-1} // pkt_{1i-1} is cached copy

Transmit data obtained by encoding in previous step

}

ElseIf(data(pkt_{2i}),data(pkt_{2i-1}) !=0)

{

Perform network coding on pkt_{2i} and pkt_{2i-1}

Transmit data obtained by encoding in previous step

}

Else

```
    {
    Select either of the packet and transmit
    }
}
Else
{
    Perform network coding on pkt1 and pkt2
    Transmit data obtained by encoding in previous step
}
//End of Algorithm
Function data(packet pkt)
{
    Return data encapsulated in packet "pkt"
}
```

Algo1: Algorithm for aggregate node

4 SIMULATION RESULTS

Given algorithm is implemented in ns-2.34. Traffic is generated using a CBR traffic generator at leaf nodes. We have simulated the discussed algorithm on topologies of 20, 44 and 92 nodes. AS mentioned earlier, function ***f*** is simulated with the help of random variable. Caching is only implemented at aggregate node with cache buffer size of two, one for each leaf node. Results of simulations are given in Fig 3 for both types of network that is to say network without coding and network with coding. Each simulation is run four times and so each bar of Fig 3 represents average of four simulation runs. This is done to mitigate the effect of random variable and simulation parameters.

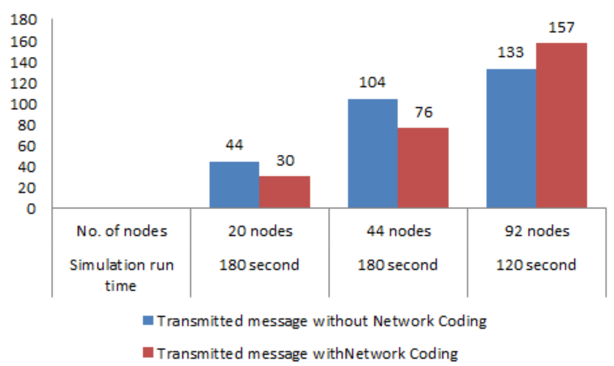


Fig3: Simulation result

One could easily see that for smaller network size (20, 44) total numbers of transmissions per unit time have reduced. But for the network of 92 nodes it has actually increased. This could possibly be because for large networks, numbers of collisions in network, without coding are more as compared to one with network coding.

5 CONCLUSIONS AND FUTURE WORK

In this paper we have implemented the idea presented in [4]. We have found out that as the size of sensor network increases, approach with network coding allows better bandwidth utilization. Though we have not quantified, one could easily argue, as number of transmissions required to send one packet from leaf node to sink nodes decreases it also provides significant energy savings at sensor nodes. Although the topology suggested is scalable and robust due to multiple paths from leaf to sink nodes, the cost effectiveness of this topology still remains an open question. It is also challenging to come up with a good ***f***, as it depends a lot upon the application and environment.

6 REFERENCES

[1] Ian F. Akyildiz, Weilian Su, YogeshSankarasubramaniam, and ErdalCayirci, 2002, A Survey on Sensor Networks, Georgia Institute of Technology.

[2] Rudolf Ahlswede, NingCai, Shuo-Yen Robert Li, Raymond W. Yeung, 2000, Network Information Flow, IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 46, NO. 4.

[3] Young-ri Choi , Mohamed G. Gouda , UTCS Technical Report TR-05-03 , Disciplined Flood Protocols in Sensor Networks, The University of Texas at Austin.

[4] R. R. Rout , S.K. Ghosh , S. Chakrabarti,2009, Fourth International Conference on Industrial and Information Systems, ICIIS 2009, 28 - 31 December 2009, Sri Lanka . Network Coding-aware Data Aggregation for a Distributed Wireless Sensor Network, Indian Institute of Technology, Kharagpur.

[5] P. A. Chou, Y. Wu, and K. Jain, 2003, Practical network coding, in Proc. Allerton Conf. Commun.,

[6] SudiptaSengupta, ShravanRayanchu, and Suman Banerjee, Network Coding-Aware Routing in Wireless Networks.

[7] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, 2006, Proc. ACM SIGCOMM, Sep. 2006, pp. 243–254 , XORs in the air: Practical wireless network coding.

[8] Qunfeng Dong , Jianming Wu ,Wenjun Hu, Jon Crowcroft ,Practical Network Coding in Wireless Networks ,University of Wisconsin & Cambridge.

[9] Raymond W. Yeung, Shuo-Yen Robert Li, NingCai, Zhen Zhang, Network Coding Theory.

[10] K V Rashmi, Nihar B Shah and P Vijay Kumar, 2010, Network Coding, IISc.

[12] AnanthKini, Vilas Veeraraghavan, Steven Weber,2004,Fast and efficient randomized flooding on lattice sensor networks, Department of Electrical and Computer Engineering Drexel University.

[13] Ting Zhu, ZiguoZhong, Tian He, and Zhi-Li Zhang, Exploring Link Correlation for Efficient Flooding in Wireless Sensor Networks, University of Minnesota.

[14] Christina Fragouli, J¨orgWidmer, Jean-Yves Le Boudec,On the Benefits of Network Coding for Wireless Applications.