International Journal of Latest Trends in Engineering and Technology
Vol.(10)Issue(1), pp.184-189
DOI: http://dx.doi.org/10.21172/1.101.33
e-ISSN:2278-621X

OPTIMIZATION OF TRANSMISSION EFFORTS IN FAULTY NETWORKS VIA NETWORK CODING

Jaskarn Singh Bhullar[1], Sonia Sharma[2]

Abstract-Network coding is vulnerable to pollution attacks when malicious nodes injects some bogus data blocks or modifies the contents of the packets. When these data streams are combined with other legal blocks of the network at downstream nodes, it makes decoding of the original blocks cumbersome and ultimately degrades the network performance. In wiretapping, some adversary can wiretap the communication channels and attains the control of information for viewing purposes, while in byzantine attack the attackers can also modify the coded packets along with the information viewing. To address this issue, a novel approach to limit pollution attack by selecting the optimal network path for packet transmission is specified in this work. This scheme rapidly identifies the malicious node of the network and marks it inaccessible, so that the system can quickly recover from pollution attack and can disseminate data to the

1. INTRODUCTION

In overlay networks, the upstream nodes broadcast the data packets to its downstream for end to end reliable data transmission. But these networks are more vulnerable to the pollution attacks from these cooperative end hosts. During the transmission an attacker can also gain the control of the link/receiver and lead to the loss of data privacy. Wiretapping (Type I) and Byzantine modification (Type II) are the two main types of attacks in network coding environment In wiretapping, some adversary can gain the control of the communication link for information viewing purposes. Although the data can't be altered but the private information can be broadcasted to the other users. Byzantine attack means that the attackers can also modify the coded packets along with the information viewing. In this case, the altered packets may confuse the receiver and force to take some wrong decisions. Modification detection of source packets is also very important as compared with the modification correction.

In this work, an attempt has been made to secure the network coding[1] by combining information-theoretic approaches[2] with cryptographic approaches. This scheme has high probability of detecting the modifications because each source packet is augmented with a hash symbol which is computed from a simple nonlinear polynomial function of the data symbols.

2. RELATED WORK

Various schemes for handling network-coding pollution attacks are divided into three main categories as 1) Attack detection 2) Identification of attacker 3) Error correction. The first category allows intermediate nodes to verify the blocks on the fly. Typically, these schemes are based on public key crypto-systems having homo-morphic properties, such as homomorphic hashing[3] and homomorphic signatures[4]. These schemes require expensive exponentiation computations at each hop, thus incurring large computational delays. To moderate computational costs, probabilistic checking[5] or null space properties[6] of network coding has been proposed. However in these schemes, the verification information is derived from the blocks to be propagated, so it must be repeatedly pre-distributed to all the nodes. These schemes also lead to significant delays and communication overheads.

The approaches that are used for identifying the pollution attackers, introduces significant communication overhead and make the process vulnerable to colluding attackers. The schemes in third category aim at correcting corrupted blocks at the sink nodes by introducing a level of redundancy. But error correction is applicable only when the corrupted blocks are limited and the achievable flow rate is determined by the number of contaminated links.

A model to secure the network coding and information security was proposed first time by Nakamura and Kodama[7] and the sufficient conditions against wiretapping and Byzantine modifications were laid down by Siavoshani[8]. In conventional routing, generally some encryption techniques are employed to protect against wiretapping. But the proposed work[7] showed that by only applying network coding one can securely transmit the data without incorporating cryptographic approaches. In this model message can be sent to the receiver without leaking any information to the intruder with the transmission rate of one unit. In [9], the authors showed that the random codes can also achieve the same security by using a much smaller base field than defined in [7] by scarifying a small amount of overall network capacity. Furthermore, it is

[1] Associate Professor, Department of Applied Sciences, Malout Institute of Management & Information Technology, (MIMIT) Malout, Punjab, India.

[2] Associate Professor, Department of Information Technology, Malout Institute of Management & Information Technology (MIMIT) Malout, Punjab, India.

pointed out that a large field size may sometimes be required to achieve security without giving up any capacity. Few models based on cryptographic approaches[10], signature based schemes[11][12] and polynomial hashing [13] in random network coding have also been proposed for error correction.

3. ADVERSARY MODEL

Adversary is considered as an internal or external wire tapper with computation bounded power, aims at intercepting packets and decoding to extract meaningful information. Without loss of generality, it is assumed that the source and sinks are always trusted and can never be compromised by an adversary.

3.1 Description of the scheme

When the network contains malicious nodes and faulty links, the end to end data delivery is ensured by steering the packets through the dynamic selection of the shortest route only. The faulty nodes encountered in the path are removed immediately to minimize the communication overhead. For the entire path, the total transmission time has been calculated with traditional routing and with network coding methods, to calculate the transmission time coding gain. The proposed scheme is a block-based scheme which possesses high computational efficiency and minimizes computational delays.

3.2 Algorithm for Optimal Node Selection

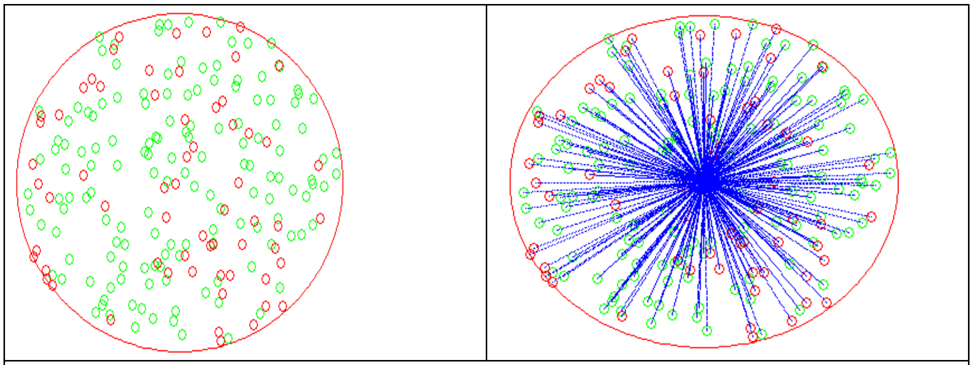
```
Problem description: Find the optimal node over the network
Input: Number of nodes: n
Output: Returns the Optimal node from the given source node
Step 1: Initialize source node, Let Source node=1.
Step 2: Loop            // To find the optimal node
    For i=2 to n
     a)  Calculate the distance from the source node (Xa,Ya)  to its fellow node (Xc,Yc) by using the
         following Euclidian’s formula:
        Di(a,c)= $\sqrt{(Xa-Xc)^2-(Ya-Yc)^2}$
     b)  If (Di<=2) then
         i)      Enter the neighbor node No.(#i) in the matrix ‘M’ defined above.
         ii)     Now Calculate the distance of #i from the sink node i.e. Dist(c,s)).
         iii)    Enter (Dist(c,s)) in Matrix M.
        Else
                  Continue
         endIf
     endFor
         From M, choose the node having minimum (dist(c, s)) as the optimal node.
  Step 3: Return optimal node.
```

4. SIMULATION AND ANALYSIS

In this part the proposed scheme has been evaluated for calculating the total transmission time taken through the dynamic networks

4.1 Analysis of Average Malicious Node Identification Time (AMNIT)

The proposed scheme is simulated in MATLAB 2014b with varying network sizes of radius (5-30) meters and error probabilities from (5-50%) to calculate the AMNIT. One sampled network of radius 10m and error probability as 30 % is shown in figure 1.



Network Size = 10 ; Error Probability = 30%

Figure 1: A sample network of radius = 10 m and error probability= 30% for calculating the Average Malicious node

Identification Time (MNIT)
Table 1 summarizes the AMNIT with varying network sizes over 6 independent runs of the same network and indicates that with the increase of network size and error probability, the total time to identify faulty nodes is also increasing but the mean time for malicious node identification is same i.e 1.96 Seconds

Table 1: Mean Malicious Node identification time with varying Network size radius (5-25) and Error probabilities (5-30) Percent

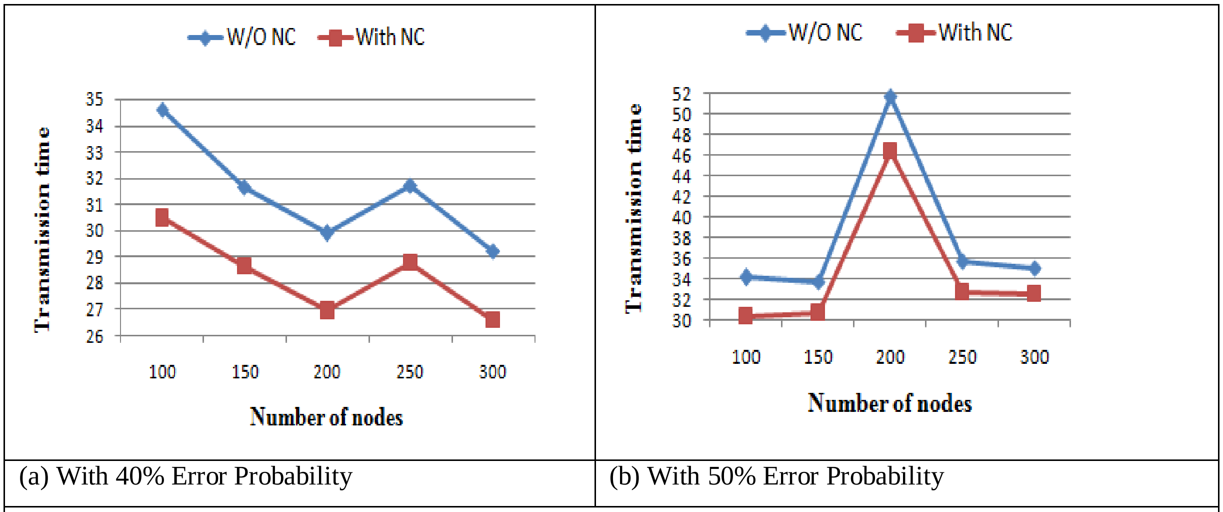
Network Radius	Total number of Faulty Nodes Encountered	Average % of Faulty Nodes Identified	Average No. of Faulty Nodes	Average Malicious Node Identification Time (Sec)
5	59	95%	56	1.899
7	111	94%	104	1.945
10	222	97%	215	1.987
15	489	98%	480	1.95
20	861	99%	853	1.979
25	1340	99%	1328	1.98
Mean Malicious Node Identification Time				1.96 Sec.

4.2 Analysis of Average Transmission Time

The total transmission time can be analyzed over a dynamic network with parameters like size of the network, error probability, number of senders and receivers, total number of bits transmitted and the total hop distance between sender and sink nodes.

Table 2: Comparison of total transmission time for a Dynamic network having 100 node with varying (0-50) % Error probability

Number Error Probability	Without network coding		With network coding		Transmission Time Coding Gain = $\frac{(\omega_a-\omega_b)}{\omega_a}$
	Total Number of Hops Visited	Total Transmission Time in Sec (ω_a)	Total Number of Hops Visited	Total Transmission Time in Sec (ω_b)	
0%	10	24.607	8	19.658	0.20
10%	9	23.748	8	19.2	0.19
20%	9	23.852	8	19.909	0.17
30%	10	26.609	9	22.897	0.14
40%	12	34.592	11	30.273	0.12
50%	12	34.105	11	30.347	0.11



(a) With 40% Error Probability

(b) With 50% Error Probability

Figure 2 : Graphical representation of Total Transmission time with the change of Network size at (a) 40% and (b) 50% error probabilities

The figures 2 shows that the transmission time for the network coding scenario is lesser than the traditional routing, for different network sizes with varying error probabilities. Moreover the transmission time increases with the increase of network size and error probability.

4.3 Analysis of Transmission Time Coding Gain

The Coding gain (CG) is the ratio of gain in the transmission to the initial value. If ω_a is the transmission time without network coding and ω_b is transmission time with coding.

Then Coding Gain (CG) = $\frac{(\omega_a - \omega_b)}{\omega_a}$

Coding gain for the variable sized Networks (100-300) nodes in the intervals of 50 nodes under varying error probabilities (0-50) for total transmission time is summarized in Table 3.

Table 3: Comparison of Transmission time Coding Gain of the proposed scheme for the varying network sizes and error probabilities

Sno.	Network Nodes	Percentage Error probability for the Transmission Time Coding Gain					
		0%	10%	20%	30%	40%	50%
1	100	0.20	0.19	0.17	0.14	0.12	0.11
2	150	0.21	0.18	0.16	0.11	0.10	0.09
3	200	0.18	0.17	0.15	0.13	0.10	0.10
4	250	0.14	0.13	0.11	0.10	0.09	0.08
5	300	0.13	0.12	0.10	0.10	0.09	0.07

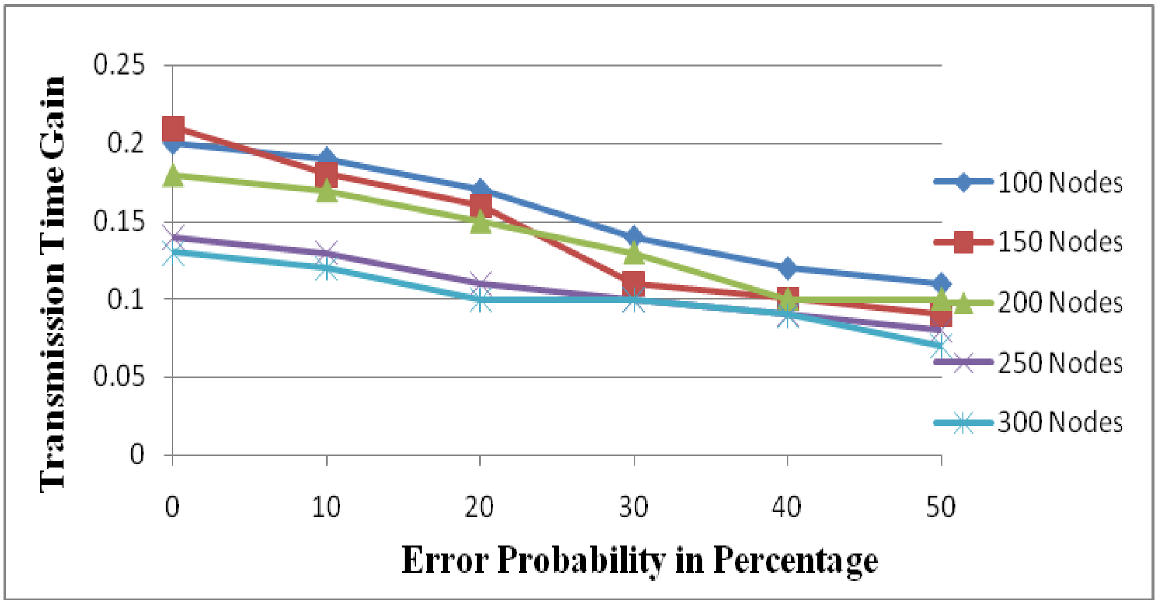


Figure 3: Graphical representation of Transmission time coding gain for varying error probabilities (0-50) in the interval of 10% and network size (100-350) in the span of 50 nodes.

Table 3 indicates that as the size of the network increases, the coding gain decreases. It has also been analyzed that when the error probability increases, the transmission time coding gain is further decreases. The maximum transmission time gain achieved with the scheme is 21% and the minimum gain is 7%

4.4 Comparison with Other Previously Existing Schemes

The performance of IMN is also compared against some already existing schemes like Homomorphic Hash[3], Trapdoor Hash[14], Mac based[15][16],Homomorphic Signature[4] and Null Keys[6] in terms of the computational Efficiency and is summarized in the table 4.

Table 4: Comparison of Transmission time Coding Gain of Various Schemes on the basis of malicious node identification time

Sno.	Schemes	% Error probability for Transmission Time Coding Gain						
		0%	10%	20%	30%	40%	50%	60%
1	HOMOMORPHIC HASHING [3]	0.111	0.102	0.088	0.077	0.077	0.070	0.062
2	HOMOMORPHIC SIGNATURE [4]	0.136	0.128	0.116	0.105	0.111	0.097	0.089
3	MAC-BASED [15]	0.151	0.143	0.129	0.118	0.123	0.110	0.100
4	NULL KEYS[6]	0.155	0.147	0.133	0.124	0.127	0.118	0.104
5	MIS[17][14]	0.082	0.077	0.068	0.061	0.066	0.058	0.051
6	IMN	0.237	0.219	0.189	0.163	0.151	0.128	0.118

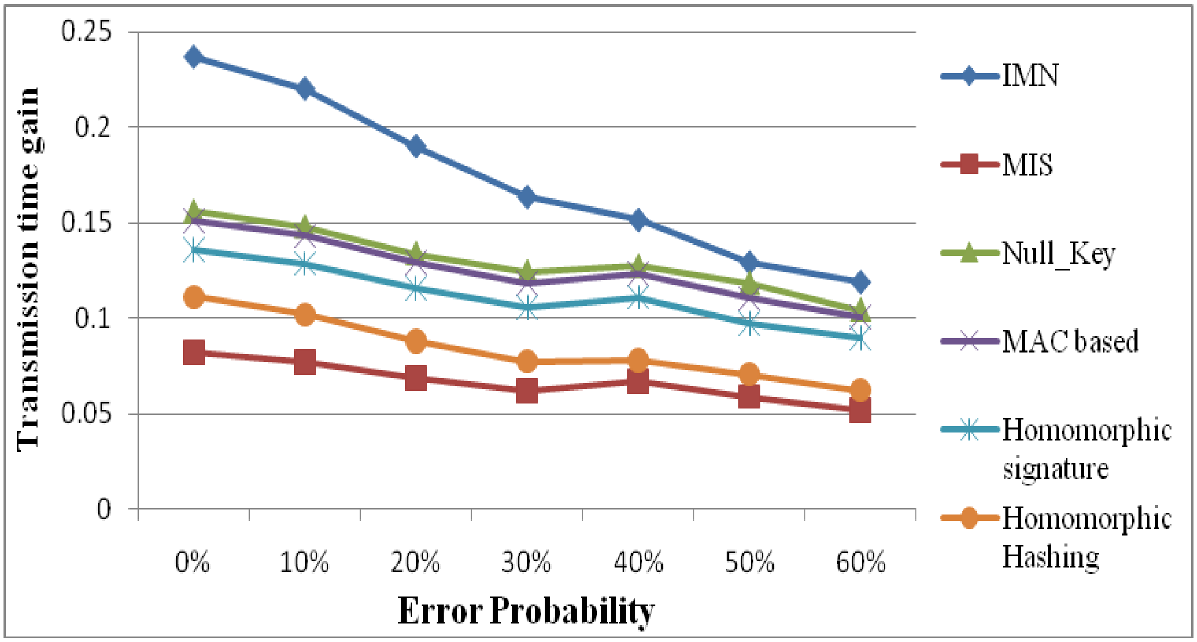


Figure 4 : Graphical representation of Transmission time coding Gain of various schemes with varying error probabilities (0-60) percent in the interval of 10%.

It has been analyzed that for the sampled network the proposed scheme (IMN) performs better than the other previous schemes and improves the network efficiency. It has also been analyzed that with the increase of error probability, the coding gain decreases. When this scheme is applied to the larger block sizes, the average data transmission time is further decreases and the network throughput increases.

5. CONCLUSION

For the data dissemination from source to the sink, a series of intermediate nodes have been encountered. The selection of an optimal path for packet forwarding is based on the choice of the optimal node. The proposed scheme quickly discovers the malicious node and declares it faulty. For the reliable transmission, the total transmission time of the network is analyzed with the traditional routing and with Network coding under varying network sizes and error probabilities. The computational efficiency of IMN is very high and it encounters very small space overhead. Furthermore, IMN does not require repeatedly distributing verification information and reduces the communication overheads. The proposed scheme can identify up to 96% of the malicious nodes of the network. The maximum transmission time gain achieved with IMN is 21% and the minimum gain is 7%. It has also been analyzed the network performs better at lower error probabilities for transmission time coding gain as compared to higher rates of malicious nodes.

6. REFERENCES

[1] Bassoli, R., Marques, H., Rodriguez, J., Tafazolli, R., (2013), “Network Coding Theory: A Survey,” IEEE Communications Surveys & Tutorials, Vol. 15, Issue. 4, pp. 1950-1978.

[2] Azari, L., Ghaffari, A., (2016), “Proposing a Novel Method based on Network Coding for Optimizing Error Recovery in Wireless Sensor Networks,” Indian Journal of Science and Technology, Vol. 8 No. 9, pp. 859–867.

[3] Cai, N., and Yeung, R.W., (2011), “Secure network coding on a wiretap network,” Information Theory, IEEE Transactions Vol. 57, No. 1, pp. 424-435.

[4] Jiang, Y., Zhu, H., Shi, M., Shen, X., and Lin, C., (2010), “An Efficient Dynamic-Identity Based Signature Scheme for Secure Network Coding,” Computer Networks: The International Journal on Computer and Telecomm. Networking, Vol. 54, No. 1, pp. 28-40.

[5] Yao, H., Silva, D., Jaggi, S., and Langberg, M., (2010), “Network codes resilient to jamming and eavesdropping,” in IEEE Int. Symp. Network Coding, Available: http://arxiv.org/abs/1001.3714

[6] Wang, J., Wang, J., Zhu, Y., and Jia, C., (2015), “An Efficient Short Null Keys Based Scheme for Securing Network Coding Against Pollution Attacks” Springer-Verlag Berlin Heidelberg, CCIS 502, pp. 16–31.

[7] Nakamura, M., and Kodama, T., (1990), “An efficient hybrid ARQ scheme for broadcast data transmission systems,” IEICE Transactions, Vol. J73-A, pp. 277–283.

[8] Siavoshani, M. J., Fragouli, C., and Diggavi, S. N., (2008), “On Locating Byzantine Attackers,” in Network Coding Workshop on Theory and Applications.

[9] Feldman, J., Malkin, T., Servedio R. A., Stein, C., (2004), “On the capacity of secure network coding,” Proceedings of the 42nd Annual Allerton Conference on Communication, Control, and Computing.

[10] Johnson, R., Molnar, D., Song, D., and Wagner, D., (2002), “Homomorphic Signature Schemes,” RSA Conference-Cryptographers.

[11] Gkantsidis, C., and Rodriguez, P., (2005), “Network Coding for Large Scale Content Distribution,” IEEE Infocom.

[12] Boneh, D., Goh, E. J., Lynn, B., and Shacham, H., (2003), “Aggregate and Verifiably Encrypted Signatures from Bilinear Maps,” Advances in Cryptology — Eurocrypt.

[13] Ho, T., Leong, B., Koetter, R., Medard, M., Effros, M., Karger, D. R., (2008), “Byzantine modification detection in multicast networks with random network coding,” IEEE Transactions on Information Theory, Vol. 54, No. 6, pp. 2798-2803.

[14] Yang, F., (2015), “Efficient Trapdoor Hash Functions for Digital Signatures,” International conference on soft computing and Software engineering (SCSE 2015), pp. 351-357.

[15] Shaobin, C., Zhenguo, G., DeSen, Y., Nianmin, Y., (2013), “A network coding based protocol for reliable data transfer in underwater acoustic sensor,” Ad Hoc Networks 2013, Vol. 11 No. 5 pp. 1603–1609.

[16] Yutao, L., Benun, V., and Jianbin, W.,(2012), "A Novel Acoess Selecti on Scheme in Heterogeneous Wireless Environm ents", IJACT VoL. 4 No. 1 pp. 24-32.

[17] Wang, Q., Vu, L., Nahrstedt, K., and Khurana, H., (2010), “Identifying Malicious Node in Network-Coding-Based Peer-to-Peer Streaming Networks,” in Proc. IEEE Mini INFOCOM 2010 San Diego, CA, USA, pp. 1-5.