IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 1, March, 2013
ISSN: 2320 - 8791
www.ijreat.org

Enhancing the Security of Content Distribution using On-the-Fly Verification via Network Coding

Abarna.R[1], Krupahari.A.L[2] and PraveenKumar.R[3]

[1,2,3] Information Technology, Anand Institute of Higher Technology, Chennai.

Abstract

Content distribution in network may be vulnerable henceforth unauthorized users can inject "bogus" data to corrupt the content during its distribution process in order to deplete the network resource. Content verification is the important one to maintain integrity of the content. This is achieved by on the fly verification. Content is splitted and hashing is applied. Hashed content is sent to the destination through peer to peer network. Key is sent to the destination directly from the source. Without the key unauthorized users finds difficult to modify the content.

Keywords: *Content Distribution, bogus data, network coding, verification, peer to peer network*

1. Introduction

Network Security is used to provide security to the authorized data which is being distributed from source to destination in the network. It also prevents unauthorized access of data by developing a secure network using security services like access, confidentiality, authentication, integrity, non-repudiation. Some of common internet attack methods used to modify the authorized data are eavesdropping, viruses, worms, Trojan, IP spoofing, denial of service. To prevent data from such attacks we use technologies of cryptographic systems, firewall, intrusion detection systems, anti malware software and Scanners and secure socket layer. Current development in the network security is the biometrics and smartcard which

greatly reduces the unauthorized access of secure systems. Even though we provide high security to the data there is still possible of hacking the data. Thus achieving 100% security in the network is not possible.

In existing, content is sent from the source to destination. By applying hash technique we get hashed content and key. Same key is used for the same content every time. Hashed content is sent to the destination through the centralized server . If the destination does not have the capacity of storing the content which is received from the source then both transmission rate and delay will be too high. Thus by sending same key for the same content unauthorized users easily modify the content.

In proposed, we use three techniques to maintain integrity of the content being distributed from the source to the destination. First we use Random linear network coding [1][2] which is used to split the content and to store the content in different storage locations randomly. Thus by splitting the content the transmission rate and delay will be less and network traffic will also be avoided. Homomorphic hash function [3][4] is the second technique to hash the splitted content and to generate the keys randomly. Hashed content is randomly distributed to the distributed network. Hashed content strength and keys is stored in the source. Finally we use on the-fly-verification in which we use three methods. Data verification[10] verifies hashed content and hashed content strength. Block by block downloading downloads hashed content from distributed network, random keys and hashed content strength from the source. Using keys we dehash the hashed content to get original splitted content Reconstruction reconstructs the splitted

content which is received randomly to get the original content sent from the source.

2. Definitions

Content distribution is the process of transmitting the messages or data from source to destination. Content Distribution is the act of sharing or circulating content with other websites, directories, or users. Content Distribution is a great means for product companies to circulate their products through various online means.

Network coding [6][7] is the set of techniques or algorithm for giving security during transmission via networks. Network coding is a technique which can be used to improve a network's throughput, efficiency and scalability, as well as resilience to attacks and eavesdropping, as compared to traditional methods of OSI model or TCP/IP model.

Bogus data is to insert fake data to the original data by unauthorized users.

Peer to Peer network [8][9] is also known as distributed network that interconnects number of systems within the network. It is defined as one computer in the network can act as a client or server for other computers in the network allowing shared access files and other resources such as peripherals and sensors without the need of central servers.

Content verification [10] means verifying the contents with its strength to check whether the received content is modified by unauthorized users.

3. Literature review

April 2005, S.Acedanski, S.Deb, M.Medard, and R.Koetter, Multiple storage locations are available but limited space is consumed. Each storage location chooses a part of the file without the knowledge of what is stored in the other locations. The problem is storing a large file in a distributed manner over a network.

1994, M. Bellare, O. Goldreich, and S. Goldwasser, Once we applied the transformation to some document M, the time to update the result upon modification of M should be "proportional" to the "amount of modification" done to M.

May 2004, M.N. Krohn, M.J. Freedman, and D.Mazieres, The quality of peer-to-peer content distribution can suffer when malicious participants intentionally corrupt content. Using simple block-by-block downloading we can verify blocks with signatures and hashes, but not useful when we use rate less erasure codes.

1998, M. Bellare, J. Garay, and T. Rabin, Many tasks in cryptography call for verification of a basic operation like modular exponentiation is simple and slow.

2005, M.Wang, Y.Zhu, B.Li, Large Volume of data in the overlay network seeks to design and implement the best strategy to disseminate data.

4. Proposed Model

In Fig.1, We are using three algorithms Random Linear Network Coding [1][2], Homomorphic Hash Function [3][4], On the Fly Verification.

4.1 Random Linear Network Coding

In which the original content is splitted and stored in different storage locations randomly. Each storage locations does not know what is stored in other locations and where the other part of the content is stored.

4.2 Homomorphic Hash Function

In which splitted content is hashed and then hashed content and key will be generated

automatically. Hashed content is sent to the peer to peer network. Generated key and hashed content strength is kept in the source.

4.3 On the Fly Verification

In which three methods is followed, Data Verification [10], Block by Block Downloading, Reconstruction to maintain the high level integrity of the content. In data verification hashed content and hashed content strength is verified to check whether the content is modified during distribution by unauthorized users. Only if it is not modified hashed content and key is downloaded by Block by Block Downloading. Once the hashed content and key is received, hashed content is dehashed by using keys. Now to get the original content, content which is distributed randomly is reconstructed.

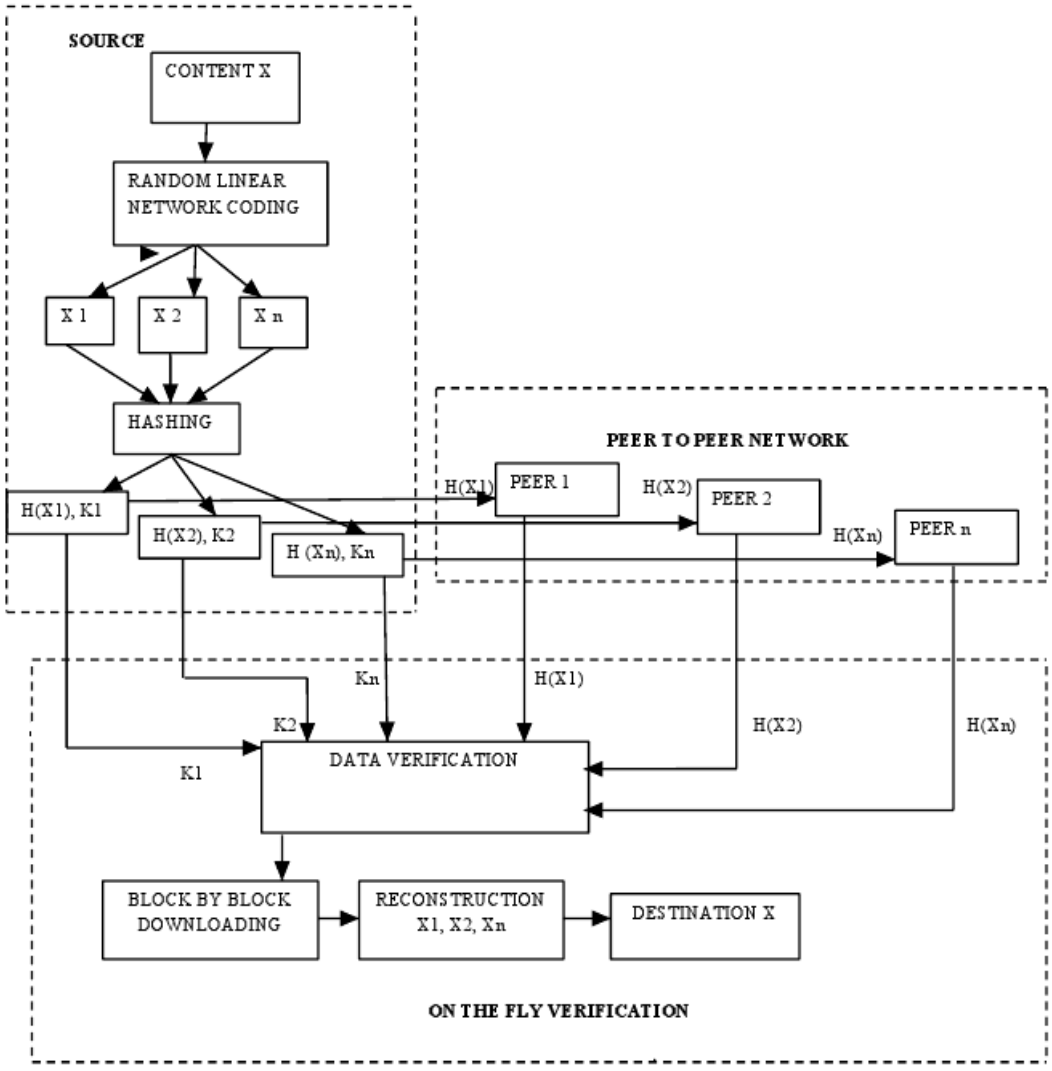

Fig.1 On the Fly Verification

Step 1: In the source original content is splitted into n parts by using Random Linear Network Coding.

Step 2: Hashing is applied to splitted content. Hashed content and key will be generated.

Step 3: Hashed content is sent to peer to peer network.

Step 4: Hashed content and hashed content strength is verified by using data verification.

Step 5: Hashed content and keys are downloaded using Block by block Downloading.

Step 6: Using keys hashed content is dehashed to get splitted content.

Step 7: Splitted content is reconstructed because it is received randomly.

Step 8: Now the original content is received by the destination.

5. Conclusion

Network coding is to improve system throughput or peer to peer networks to improve overall system efficiency. In proposed we investigate security and efficiency issues in large content distribution based on network coding. Our proposed techniques on the fly verification based on faster homomorphic hash function have proved its security and also reduced network traffic, and delay.

6. References

[1] S. Acedanski, S. Deb, M. Medard, and R. Koetter, "How Good Is Random Linear Coding Based Distributed Networked Storage," Proc. Workshop Network Coding, Theory and Applications, Apr. 2005.

[2] S.R. Li, R.W. Yeung, and N. Cai, "Linear Network Coding," IEEE Trans. Information Theory, vol. 49, no. 2, pp.371-381,Feb 2003.

[3] M.Bellare, O.Goldreich, and S.Goldwasser, "Incremental cryptography: the case of hashing and signing", Proc. CRYPTO 1994.

[4] S.Contini, A.K. Lenstra, and R.Steinfeld, "VSH, an efficient and provable collision-resistant hash function" , Proc. EUROCRYPT, pp.165-182,2006.

[5] M.N.Krohn, M.J. Freedman and D.Mazieres,"On the fly verification of rate less erasure codes for efficient content distribution", Proc. IEEE Symp.Security and privacy, pp.226-240, May 2004.

[6] C.Gkantsidis and P.R. Roudriguez," Network coding for large scale content distribution", Proc IEEE INFOCM, pp.2235-2245, 2005.

[7] M. Wang, Z. Li, and B. Li, "A High-Throughput Overlay Multicast Infrastructure with Network Coding," Proc. Int'l Workshop Quality of Service (IWQoS), 2005.

[8]C.Gkantsidis, J.Miller, and P. Rodriguez,"Anatomy of a P2P content distribution system with network coding"Proc. Int'l workshop Peer-to-peer Systems, Feb. 2006.

[9] R.T.B.Ma, S.C.M. Lee, J.C.S. Lui, and D.K.Y. Yau,"Incentive and service differentiation in P2P networks: A Game Theoretic Approach", IEEE/ACM Trans. Networking, vol.14, no.5, pp.978-991, Oct. 2006.

[10]M.Bellare, J. Garay, and T.Rabin, "Fast batch Verification for modular exponentiation and digital signatures," Proc. EUROCRYPT, 1998.