ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

Large Scale of Content Distribution Using Homomorphic Hash Function

Sridevi K[1], Senthil Nathan K[2]

M.E, Department of CSE, PSV College of Engineering and Technology, Krishnagiri, Tamilnadu, India[1]

Assistant Professor, Department of CSE, PSV College of Engineering and Technology, Krishnagiri, Tamilnadu, India[2]

Abstract: Privacy threat is one of the critical issues in wireless networks, where attacks such as traffic analysis and flow tracing can be easily launched by a malicious adversary due to the open wireless medium. Network coding has the potential to thwart these attacks since the coding/mixing operation is encouraged at intermediate nodes. However, the simple deployment of network coding cannot achieve the goal once enough packets are collected by the adversaries. On the other hand, the coding/mixing nature precludes the feasibility of employing the existing privacy-preserving technique. In this paper, we propose a novel network coding based privacy-preserving scheme against traffic analysis in wireless networks. In Homomorphic encryption, the proposed scheme offers two significant privacy-preserving features, packet flow untraceability and message content confidentiality, for efficiently thwarting the traffic analysis attacks. Moreover, the proposed scheme keeps the random coding feature. Theoretical analysis and simulative evaluation demonstrate the validity and efficiency of the proposed scheme.

Keywords: Content distribution, security, verification, network coding

I. INTRODUCTION

In our system the main objective of the system design is to reduce the computational cost and also provide efficient security in the network. Initially the network has to be set up. For that the IP address is needed, it can be determined by giving a command 'net view' in the command prompt which will display the system name of the all other system. Then by command 'ping system name' I will get the IP address of the particular system. The senders' IP address will be given to the all other routers for establishing a network. Following by this the sender will browse the file from the system database which is to be transmitted. For distributing in each router, the file is splitted by dividing the total file size into 1000 bits by which 'n' number of packets is generated for transmission. The routers are selected through which the packet has to be sent by the sender. The password which is user defined is assigned for the verification of the client in the receiver end. This password is stored in the database for later use. The packet which is to be sent is encrypted by generation of the key using RSA algorithm. Then the packet is encrypted to which the digital signature is appended to maintain the integrity of the data in packets. The packets are decrypted and assembled in the receiver end but for assembling the packets the client has to give the password. If and only if the password is correct then only the client can get the file. The password is checked internally in the database with the password given by the sender. If the client enters a wrong password then he/she cannot get the file. The main thing which is implemented in our system is that if there is any presence of hacker in the router it will be shown in the receiver end. The client can just check in which router the hacker has come but the hacker cannot inject any bogus packets. The digital signature is generated by DSA algorithm. The combinations of these packets are forwarded to the selected routers.

II. ABOUT THE SYSTEM

This network security is used in our proposed system which will provide security to the packets while transmitting in the network established. In network security system usually consists of many components. Ideally, all components work together, which minimizes maintenance and improves security. Effective network security targets a variety of threats and stops them from entering or spreading on your network. RSA is an encryption and authentication system that uses an algorithm developed in 1977 by Ron Rivets, Adi Shamir, and Leonard Adelman. The RSA algorithm is the most commonly used encryption and authentication algorithm. The algorithm involves multiplying two large prime numbers (a prime number is a number divisible only by that number and 1) and through additional operations deriving a set of two numbers that constitutes the public key and another set that is the private key. Once the

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

keys have been developed, the original prime numbers are no longer important and can be discarded. Both the public and the private keys are needed for encryption /decryption but only the owner of a private key ever needs to know it. The private key is used to decrypt text that has been encrypted with the public key. A digital signature uniquely identifies the originator of digitally signed data and also ensures the integrity of the signed data against tampering or corruption. For creating digital signatures is the Digital Signature Security Standard (DSS) that was developed by the National Security Agency and adopted by the United States government as its digital-signature standard. DSS defines the Digital Signature Algorithm (DSA). DSA is used only for digital signatures and makes no provisions for data encryption. Only a digital signature created by the appropriate private key decrypts and validates properly with the public key. If a different private key was used to sign the data, the validity check fails. If the contents of digitally signed data or the digital signature have been tampered with or are corrupted, the validity check also fails. Valid digital signatures can be used to ensure the integrity of digital data against corruption.

III. SYSTEM ARCHITECTURE DESIGN

The system analysis is to determine where the problem is in an attempt to fix the system. The step involves the breaking down the system in different pieces to analyze the situation, analyzing project goals, breaking down what needs to be created and attempting to engage users so that definite requirements can be defined. The system architecture of the proposed system is shown in the Figure 1 which explains about securing the data in a wireless network. Connection is established based on the IP address available in the network. Using the IP address of the sender, the routers and the client are connected. Connection establishment is verified on the user interface. The file to be transmitted is browsed from the database by the sender and these retrieved file is splitted into small packets. The routers through which the packets are transmitted is selected by the sender. These routers are used for faster transmission of packets in the network. The user defined password is assigned so that the client can retrieve the file if and only if the password is correct. For encryption technique the RSA algorithm is used. In this encrypted packet the digital signature is appended to maintain the integrity of data in the packet and also for authenticity of sender. This combination of packet is sent to the client through routers. The decryption process takes place in the client side. The client first has to give the correct password to retrieve the file. After verification of password, if and only if the password matches the decryption process takes place internally and the packets will be assembled based on the digital signature. Then the complete file will be displayed to the client. The database which is used here stores the password for verification in the client side. This database plays an important role in finding the presence of the hacker in the router. . Connection is established based on the IP address available in the network. Then the complete file will be displayed to the client. The database which is used here stores the password for verification in the client side.

Design is the first step in the development phase for any techniques and principles for the purpose of defining a device, a process or system in sufficient detail to permit its physical realization. Once the software requirements have been analyzed and specified the software design involves three technical activities – design, coding, implementation and testing that are required to build and verify the software. The design activities are of main importance in this phase, because in this activity, decisions ultimately affecting the success of the software implementation and its ease of maintenance are made. These decisions have the final bearing upon reliability and maintainability of the system. Design is the only way to accurately translate the customer's requirements into finished software or a system. Design is the place where quality is fostered in development. Software design is a process through which requirements are translated into a representation of software. Software design is conducted in two steps. Preliminary design is concerned with the transformation of requirements into data. UML stands for Unified Modeling Language. UML is a language for specifying, visualizing and documenting the system. This is the step while developing any product after analysis. The goal from this is to produce a model of the entities involved in the project which later need to be built. The representation of the entities that are to be used in the product being developed need to be designed. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams typically are associated. These decisions have the final bearing upon reliability and maintainability of the system. Design is the place where quality is fostered in development.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

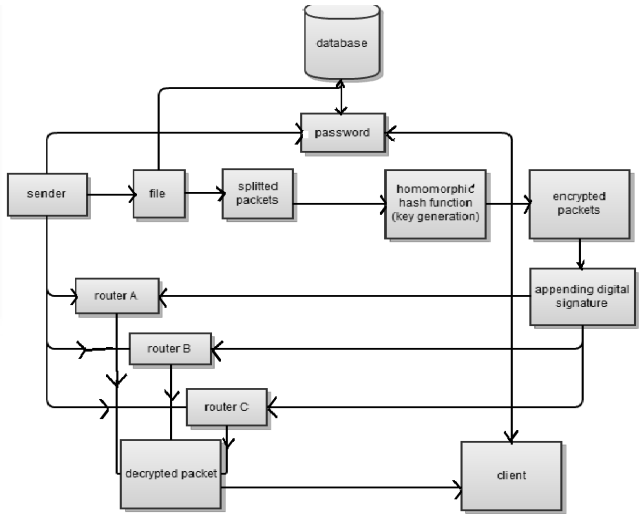

Figure 1. System Architecture Design

IV. SYSTEM ANALYSIS

4.1 RSA Encryption and Decryption Algorithm

Using an encryption key (e,n), the algorithm is as follows:

- Represent the message as an integer between 0 and (n-1). Large messages can be broken up into a number of blocks. Each block would then be represented by an integer in the same range.
- Encrypt the message by raising it to the eth power modulo n. The result is a ciphertext message C.
- To decrypt ciphertext message C, raise it to another power d modulo n

The encryption key (e,n) is made public. The decryption key (d,n) is kept private by the user.
To Determine Appropriate Values for e, d, and n

- Choose two very large (100+ digit) prime numbers. Denote these numbers as p and q.
- Set n equal to p * q.
- Choose any large integer, d, such that GCD(d, ((p-1) * (q-1))) = 1
- Find e such that e * d = 1 (mod ((p-1) * (q-1)))

4.2 Multicast with Network Coding in Application-Layer Overlay Networks

In multicast with network coding networks I seek to significantly improve end-to-end throughput in application-layer multicast by taking full advantage of these unique characteristics. This objective is achieved with two novel insights. First, i depart from the conventional view that data can only be replicated and forwarded by overlay nodes. Rather, as end systems, these overlay nodes also has the full capability To receivers needs to be a tree, and propose a novel and distributed algorithm to construct a 2-redundant multicast graph (a directed acyclic graph) as the multicast topology, on which network coding is applied. I design our algorithm such that the costs of link stress and stretch are explicitly considered as constraints and minimized. I extensively evaluate our algorithm by provable analytical and experimental results, which show that the introduction of 2-redundant multicast graph and network coding may indeed bring significant benefits, essentially doubling the end-to-end throughput in most cases. Linear codes are the coefficients that determine the linear transformations. Once a multicast graph is constructed, a set of linear codes must be found to realize linear coding multicast. I propose a distributed algorithm that is easy to implement for obtaining the linear codes for the 2-redundant multicast graph. I observe that coding is only needed at the source s and the intermediate nodes, and that an intermediate node has either one or two incoming edges (by construction).

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

Furthermore, since each leaf receiver has exactly two paths from the source s, s sends the data vector (a, b) and both a and b should be obtained by each receiver. [1]

4.3 Network Information Flow

In this paper, I study the problem with one information source, and I have obtained a simple characterization of the admissible coding rate region. Our result can be regarded as the Max-flow Min-cut Theorem for network information flow. Contrary to one's intuition, our work reveals that it is in general not optimal to regard the information to be multicast as a "fluid" which can simply be routed or replicated. Rather, by employing coding at the nodes, which I refer to as network coding, bandwidth can in general be saved. In multiterminal source coding I have proposed a new class of problems called network information flow which is inspired by computer network applications. Most classical multiterminal source coding problems in the following ways: There is no rate-distortion consideration; the sources are mutually independent; the network configuration, described by a graph, is arbitrary the reconstruction requirements are arbitrary. Our formulation covers a large class of problems instead of one particular problem most classical multiterminal source coding problems, the problem degenerates.[2]

4.4 Linear Network Coding

A communication network in which certain source nodes multicast information to other nodes on the network in the multihop fashion where every node can pass on any of its received data to others. I are interested in how fast each node can receive the complete information, or equivalently, what the information rate arriving at each node Allowing a node to encode its received data before passing it on, the question involves optimization of the multicast mechanisms at the nodes. Among the simplest coding schemes is linear coding, which regards a block of data as a vector over a certain base field and allows a node to apply a linear transformation to a vector before passing it on. I formulate this multicast problem and prove that linear coding suffices to achieve the optimum, which is the max-flow from the source to each receiving node. Network coding refers to coding at the intermediate nodes when information is multicast in a network. In this network, I want to multicast two bits b1 and b2 from the source S to both the nodes Y and Z.A solution is to let the channels SU,TW,TY carry the bit b1and channels SU,UW,UZ carry the bit b2 and channels WX,XY,XZ carry the exclusive-OR the bit b1 ⊕ b2 . The node Y receives b1 and b1⊕b2 from which the bit b2 can be decoded. Similarly, the node Z can decode the bit b1 from b2 and b1⊕b2 and the coding/decoding scheme is assumed to have been agreed upon beforehand. It is not difficult to see that the above scheme is the only solution to the problem. In other words, without network coding, it is impossible to multicast two bits per unit time from the source S to both the nodes Y and Z. This shows the advantage of network coding. In fact, replication of data can be regarded as a special case of network coding.[3]

4.5 Polynomial Time Algorithms for Multicast Network Code Construction:

The max-flow min-cut theorem states that a source node S can send information through a network (V,E) to a sink node at a rate determined by the min-cut separating s and t. Recently, it has been shown that this rate can also be achieved for multicasting to several sinks provided that the intermediate nodes are allowed to re-encode the information they receive. I demonstrate examples of networks where the achievable rates obtained by coding at intermediate nodes are arbitrarily larger than if coding is not allowed. I give deterministic polynomial time algorithms and even faster randomized algorithms for designing linear codes for directed acyclic graphs with edges of unit capacity. I extend these algorithms to integer capacities and to codes that are tolerant to edge failures. A polynomial time algorithm for centralized design of optimal network multicast codes. The codes are linear with symbols from a finite field F. In practice, i will use a field of size

$$|\mathbb{F}| = 2^m$$

So that the edges actually carry bits. Coding is done by forming linear combinations of the field elements reaching a node.[4]

4.6 Cooperative Security for Network Coding File Distribution

This paper presents a practical security scheme for network coding that reduces the cost of verifying blocks on-the-fly while efficiently preventing the propagation of malicious blocks. In our scheme, users not only cooperate to distribute the content, but (well-behaved) users also cooperate to protect themselves against malicious users by

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

informing affected nodes when a malicious block is found. I analyse and study such cooperative security scheme and introduce elegant techniques to prevent DoS attacks. I show that the loss in the efficiency caused by the attackers is limited to the effort the attackers put to corrupt the communication, which is a natural lower bound in the damage of the system. I also show experimentally that checking as low as 1-5% of the received blocks is enough to guarantee low corruption rates.[5]

V. MODULES AND DEFINITION

The main objective of the system design is to reduce the computational cost and also provide efficient security in the network. Initially the network has to be set up. For that the IP address is needed, it can be determined by giving a command 'net view' in the command prompt which will display the system name of the all other system. Then by command 'ping system name'. I will get the IP address of the particular system. The senders' IP address will be given to the all other routers for establishing a network. Following by this the sender will browse the file from the system database which is to be transmitted. For distributing in each router, the file is splitted by dividing the total file size into 1000 bits by which 'n' number of packets is generated for transmission. The routers are selected through which the packet has to be sent by the sender. The password which is user defined is assigned for the verification of the client in the receiver end. This password is stored in the database for later use. The packet which is to be sent is encrypted by generation of the key using RSA algorithm. Then the packet is encrypted to which the digital signature is appended to maintain the integrity of the data in packets. The digital signature is generated by DSA algorithm. The combinations of these packets are forwarded to the selected routers. The packets are decrypted and assembled in the receiver end but for assembling the packets the client has to give the password. If and only if the password is correct then only the client can get the file. The password is checked internally in the database with the password given by the sender. If the client enters a wrong password then he/she cannot get the file. The main thing which is implemented in our system is that if there is any presence of hacker in the router it will be shown in the receiver end. The client can just check in which router the hacker has come but the hacker cannot inject any bogus into the packets advantage of the proposed system. Thus only the presence of hacker is determined by the client. The implementation modules will be explained subsequently.

5.1 FILE SPLITTING

Initially the routers are identified in the command prompt using the command "net view". The connection between the routers and the sender is checked using the ping command. After it has been done successfully the sender has to select the file to be transmitted. The file is browsed from the system database. The path of the file is shown for the sender purpose to check whether the correct file has been selected. The size of the file in shown in bits. The total number of bits is divided into 1000bits. These divided file are made as small packets. For reducing the communication cost I use three different routers for transmission before the file reaches the client. These packets undergo encryption and hashing techniques before it is transmitted. The routers through which the packets are transmitted should be selected along with the client. For the purpose of security a user-defined password is obtained. This password is stored in the database for the verification of valid client after the transmission. The file can be obtained by the client if and only if the password matches with the one stored in the database

5.2 RSA ENCRYPTION

The file which is splitted (packets) is to be encrypted using RSA algorithm. The key is generated for the encryption of packets. Here i am using asymmetric key for the purpose of encrypting and decrypting the data in packets. The data is encrypted so that the data packet is secured and hacker cannot read the encrypted data (cipher text). Then in the receiver end the packet is decrypted and the client will get the complete file.

5.3 RANDOM LINEAR NETWORK CODING

As the encryption process gets over the digital signature will be generated for each packets, it is appended at the end of each packets. This digital signature is generated using random linear network coding. This technique is the process of obtaining the signature in different formats in each time the packet is signed. As the signature changes its format each time, the hacker will not be able to identify the correct format of the packet. This increases the integrity and confidentiality of the data packet which is transmitted. During the decryption process this digital signature is separated first and it is verified.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

5.4 ATTACK MODEL

In this model the client can determine the presence of the hacker in the router. Once the client gives the password it will be matched with the password given by the sender which is stored in the database for verification. If the hacker tries to hack in the router then that will be shown to the client in the receiver end while getting the file. The hacker cannot inject any bogus into the packet in spite of his/her presence in the router. The arrival of the hacker will be shown to the client by a pop up message and it is also updated in the database. The database consists of a router table with field name as r1, r2 and r3. The default value is 0 for all the fields in the table. If there is presence of hacker in the router then the value will be updated to 1 in that respective field. AS the default value changes to 1 the field name will be shown to the client in pop up message. The hacker cannot inject any bogus into the packet in spite of his/her presence in the router.

VI. CONCLUSION AND FUTURE WORKS

6.1 Conclusion

In the application of network coding in wireless networks to improve system throughput, or P2P networks to improve overall system efficiency. In this paper, I investigate the security and efficiency issues in large content distribution based on network coding. I consider the problem of on-the-fly verification of the integrity of the data in transit. Although a previous scheme based on homomorphism hash functions is applicable, it was mainly designed for server side coding only, and will be much less efficient when it is applied on random network coding. I propose a new on-the-fly verification scheme based on a faster holomorphic hash function, and proved its security. I also consider the computation and communication cost incurred during the content distribution process. I identify various sources of the cost, and investigate ways to eliminate or reduce the cost. In particular, I propose a sparse variant of the classical random linear network coding, where only a small constant number of blocks are combined each time. Furthermore, I discuss some possible enhancements under certain conditions of the parameters, and ways to trade off among different cost. The distribution is done one step at a time. In each step, for all the nodes that have received data from all of their upstream nodes generate their own combinations and deliver them to their down- stream nodes. This process is repeated until no further delivery is possible. Finally, I examine the data received by each node, and determine if it is sufficient for the node to reconstruct the original data. I propose a sparse variant of the classical random linear network coding, where only a small constant number of blocks are combined each time where only a small constant number of blocks are combined each time.

6.2 Future Works

This work motivates several directions for future research. First, the routers are identified in the command prompt using the command "net view". The connection between the routers and the sender is checked using the ping command. After it has been done successfully the sender has to select the file to be transmitted. The file is browsed from the system database. The path of the file is shown for the sender purpose to check whether the correct file has been selected. The size of the file in shown in bits. The total number of bits is divided into 1000bits.Second, the file which is splitted (packets) is to be encrypted using RSA algorithm. The key is generated for the encryption of packets. Here I are using asymmetric key for the purpose of encrypting and decrypting the data in packets. The data is encrypted so that the data packet is secured and hacker cannot read the encrypted data (cipher text). Then in the receiver end the packet is decrypted and the client will get the complete file. Third, as the encryption process gets over the digital signature will be generated for each packets, it is appended at the end of each packets. This digital signature is generated using random linear network coding. This technique is the process of obtaining the signature in different formats in each time the packet is signed. As the signature changes its format each time, the hacker will not be able to identify the correct format of the packet. This increases the integrity and confidentiality of the data packet which is transmitted. During the decryption process this digital signature is separated first and it is verified. Finally, the client can determine the presence of the hacker in the router. Once the client gives the password it will be matched with the password given by the sender which is stored in the database for verification. If the hacker tries to hack in the router then that will be shown to the client in the receiver end while getting the file. The hacker cannot inject any bogus into the packet inspite of his/her presence in the router. The purpose of encrypting and decrypting the data in packets. The data is encrypted so that the data packet is secured and hacker cannot read or to be encrypted or decrypted data.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

REFERENCES

[1] C. Gkantsidis and P.R. Rodriguez, "Network Coding for Large Scale Content Distribution," Proc. IEEE INFOCOM, pp. 2235-2245, 2005.
[2] 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.
[3] J. Edmonds, "Minimum Partition of a Matroid into Independent Sets," J. Research of the Nat'l Bureau of Standards, vol. 869, pp. 67-72, 1965
[4] M. Bellare, O. Gold
reich, and S. Goldwasser, "Incremental Cryptography: The Case of Hashing and Signing," Proc. CRYPTO, 1994.
[5] 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.
[6] B. Fan, J.C.S. Lui, and D.-M. Chiu, "The Design Trade-Offs of BitTorrent-Like File Sharing Protocols," IEEE/ACM Trans. Networking, vol. 17, no. 2, pp. 365-376, Apr. 2009.
[7] C. Gkantsidis and P. Rodriguez, "Cooperative Security for Network Coding File Distribution," Proc. IEEE INFOCOM, pp. 1-13, Apr. 2006.
[8] M.N. Krohn, M.J. Freedman, and D. Mazie`res, "On-the-FlyVerification of Rateless Erasure Codes for Efficient Content Distribution," Proc. IEEE Symp. Security and Privacy, pp. 226-240, May 2004.
[9] P.A. Chou, Y. Wu, and K. Jain, "Practical Network Coding," Proc. Allerton Conf. Comm., Control, and Computing, Oct. 2003.
[10] R. Ahlswede, N. Cai, S.-Y.R. Li, and R.W. Yeung, "Network Information Flow," IEEE Trans. Information Theory, vol. 46, no. 4, pp. 1204-1216, July 2000.
[11] R. Koetter and M. Me´dard, "Beyond Routing: An Algebraic Approach to Network Coding," Proc. IEEE INFOCOM, pp. 122130, 2002.
[12] S. Contini, A.K. Lenstra, and R. Steinfeld, "VSH, an Efficient and Provable Collision-Resistant Hash Function," Proc. EUROCRYPT, pp. 165-182, 2006.