Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Security concerns and countermeasures in network coding based communication systems: A survey

Vahid Nazari Talooki [a,*], Riccardo Bassoli [a], Daniel E. Lucani [b], Jonathan Rodriguez [a], Frank H.P. Fitzek [b], Hugo Marques [a], Rahim Tafazolli [c]

[a] *Instituto de Telecomunicações, Aveiro, Portugal*
[b] *Department of Electronic Systems, Aalborg University, Aalborg, Denmark*
[c] *Centre for Communication Systems Research, Surrey, UK*

ARTICLE INFO

Article history:
Received 9 September 2014
Received in revised form 23 January 2015
Accepted 14 March 2015
Available online xxxx

Keywords:
Network coding
Security attacks
Eavesdropping attacks
Byzantine attacks
Corruption attacks

ABSTRACT

This survey paper shows the state of the art in security mechanisms, where a deep review of the current research and the status of this topic is carried out. We start by introducing network coding and its variety of applications in enhancing current traditional networks. In particular, we analyze two key protocol types, namely, state-aware and stateless protocols, specifying the benefits and disadvantages of each one of them. We also present the key security assumptions of network coding (NC) systems as well as a detailed analysis of the security goals and threats, both passive and active. Current proposed security mechanisms and schemes for NC in the literature are classified too. This paper also presents a detailed taxonomy of the different NC security mechanisms and schemes reported in the literature.

1. Introduction

In coding theory, there are three main coding families: source coding, channel coding, and network coding (NC). The first aims to compress information at the source, while the second introduces redundant bits at the link layer to guarantee reliable communications. On the other hand, the third consists in a coding process that takes place at intermediate nodes in the network and at different layers of the network stack.

Network coding [1] represents a generalization of classical store-and-forward routing for information flow. This new code-and-forward paradigm considers that the source messages are algebraic entities upon which operations can be performed at the intermediate nodes. This contrasts with current state-of-the-art routing solutions, where source messages are merely routed from source to destination. The main result of [1] was the enunciation of the max-flow min-cut theorem for information flow. In particular, the authors demonstrated that the multicast capacity is achieved by applying network coding. This achievement can bring a drastic change in improving the data throughput of networks. Therefore, several studies focused on implementing practical network coding based approaches that lead network coding to its current level.

In 2003, [2] demonstrated that linear operations at the nodes were sufficient to achieve the max-flow min-cut bound: linear network coding (LNC) was defined in directed acyclic networks with single-source multicast. Side by side, [3] provided an algebraic formulation of linear network coding by using abstract algebra and algebraic geometry, and demonstrated equivalent results in the

[*] Corresponding author at: Campus Universitário de Santiago, Aveiro P-3810-193, Portugal. Tel.: +351 234377900.

E-mail addresses: vahid@av.it.pt (V.N. Talooki), bassoli@av.it.pt (R. Bassoli), del@es.aau.dk (D.E. Lucani), jonathan@av.it.pt (J. Rodriguez), ff@es.aau.dk (F.H.P. Fitzek), hugo.marques@av.it.pt (H. Marques), r.tafazolli@surrey.ac.uk (R. Tafazolli).

http://dx.doi.org/10.1016/j.comnet.2015.03.010
1389-1286/© 2015 Elsevier B.V. All rights reserved.

new framework for both acyclic and cyclic networks. This algebraic formulation opened the way to random linear network coding (RLNC) [4], a kind of network codes, in which the coefficients of linear combinations are randomly chosen over a finite field. [5] provided a first description of how to implement RLNC in practice: it analyzed both its benefits in terms of throughput and its main issues. In fact, real information is flowing asynchronously so RLNC can suffer delays and losses, and it can eventually experience congestions and link failures. So, to be practical, the RLNC should be designed with a special packet structure by taking into account new key characteristics to overcome these issues.

1.1. Secure network coding

The new way of managing information, that network coding introduces in actual networks, presents new several challenges in terms of security: processing (recoding) the received data packets from neighbors in the intermediate nodes and then forwarding them, opens a myriad of challenging security issues [6]. In fact, network coding can have either positive or negative secure aspects. In the former case, by sending linear combinations of packets and not merely source data, an adversary that is intercepting some transmissions collects information that results to be useless. On the other side, coding operations across packets can make the overall network more vulnerable against several types of attacks. The research on secure network coding has been growing by mainly investigating both Byzantine and eavesdropping attacks. In fact, protocols based on network coding present vulnerability against many threats and attacks including but not limited to impersonation, Byzantine (fabrication, modification and replay) attacks, blackhole, and eavesdropping.

Some of the first analyses on secure network coding are [7–11]. These works considered an eavesdropper seeing information transmitted on a subset of network channels in a single-source scenario. In order to study secure network coding, different models has been used. [9,10] for first time proposed a model for a collection of subsets of wiretap channels for an NC system called *wiretap network*. Each wiretapper in the model has full access to only one of these subsets; however, by applying secure linear network codes, none of the wiretapper is able to extract any information from the transmitted message. At the first analyses, shown in [12,13], the measure of security has been done in terms of either information quantities or decoding probability. Next, [14] proposed an algebraic secure criterion. Side by side, the issue of designing secure network codes has been also investigated from another point of view: first in 2003 and then in 2006, [8,15,16] described how network coding could be seen as a generalization of classical error correction. In particular, network error correction (NEC) coding has been proved to be optimum in correcting random errors, erasures and errors due to malicious nodes in the network. [17] started researching on secure network coding through NEC coding. The objective of that work was the correction of errors injected by wiretapper and the protection of source messages from wiretapping.

At the best of authors' knowledge, this is the first survey to include the most relevant literature in the diverse research areas related to security attacks and mitigation techniques in network coding based communication systems. This survey includes more than 200 references and makes a broad description of security threats and attacks in network coding based systems, reviewing the current mechanisms against these security attacks and the latest results for proving a secure network coding approach. We believe both advanced and initial researchers in this area can benefit from this survey.

Other relevant surveys on principal concepts of network coding theorem are [6,18], and the most recent [19]. Some tutorial on security attacks and threats in network coding based systems and summarizing the current mitigation techniques are [20–25]. Network coding website [26] provides several studies and papers on this filed too. Also, another useful source is [27] that includes "Bibliography on Secure Network Coding" and list of paper works in the scope of secure NC systems from 2006 to 2014.

1.2. Structure of the survey

In what follows, in Section 2 we review the fundamentals of network coding, the security assumptions in NC systems and state-aware and stateless NC protocols. Then in Section 3, the security threats and attacks in the NC based systems are studied. In Section 4 we classify the proposed security mechanisms and schemes for NC in the literature. Also, we presented security techniques taxonomy in network coding based communication systems. In Section 4 we present a timeline of these mechanism and schemes too. Finally Section 5 presents the considerations and conclusions.

2. Preliminaries

2.1. Principles of random linear network coding

As an example to show the capability and the benefits of using NC in improving network throughput, Fig. 1 shows a possible simple scenario for both traditional store and forward mechanism and network coding elegant *store-process-forward* paradigm. Here the source node S wants to multicast some packets toward two sink nodes D_1 and D_2. Each packet like $p^{Packet\ Number}_{Time\ Stamp}$ has a packet number and time stamp that shows the packet number which was assigned by source node to it and the time that packet was forwarded. For simplicity each packet (or symbol) is considered as one bit.

A traditional store and forward mechanism, would achieve maximum throughput of 1.5 bits/s but NC allows both D_1 and D_2 to achieve a rate of 2 bits/s at the same time which means more than 30% improvement in throughput for this scenario.

In general, RLNC can be designed in practice according to two main approaches, called respectively intra-session and inter-session. In the former [28–31], routers combine packets belonging to the same session. It is typically used in multicast application and in case of unpredictable

Please cite this article in press as: V.N. Talooki et al., Security concerns and countermeasures in network coding based communication systems: A survey, Comput. Netw. (2015), http://dx.doi.org/10.1016/j.comnet.2015.03.010