Optimal Transmission Topology Construction and Secure Linear Network Coding Design for Multicast with Integral Link Rates

Ruimin Zhao†, Jin Wang‡, Kejie Lu§, Xiangmao Chang†, Lingzhi Li† and Yanqin Zhu†
†Department of Computer Science and Technology, Soochow University
§Department of Electrical and Computer Engineering, University of Puerto Rico at Mayagüez
‡School of Computing Science and Technology, Nanjing University of Aeronautics and Astronautics

Abstract—Linear network coding (LNC) is a promising technology that can increase network throughput, improve transmission robustness and provide data confidentiality. In this paper, we investigate the optimal transmission topology construction and LNC design for secure multicast, i.e., the Integer Secure Multicast (ISM) problem. The ISM problem aims to find the transmission topology with integral link rates and design the secure LNC to (1) meet the requirements of weak security, (2) achieve the maximum secure transmission rate (STR) and (3) minimize transmission cost (TC) under the maximum STR guarantee. Specifically, in this paper, we firstly give theoretical analysis to show a sufficient and necessary condition that there exists a transmission topology with integral link rates and a secure LNC with STR $R$. Based on the sufficient and necessary condition, we then formulate the ISM problem into an integer linear programming (ILP) and also design an efficient transmission topology construction (TTC) algorithm to obtain a near-optimal solution of the problem. Based on Lagrangian relaxation and subgradient algorithm, the TTC algorithm is designed within polynomial computational complexity. Finally, we evaluate the performance of the proposed TTC algorithm together with a lower bound and an upper bound of the ISM problem through extensive experiments. Simulation results validate the effectiveness of the proposed TTC algorithm.

I. INTRODUCTION

Over the past fifteen years, network coding (NC) theory has brought significant developments to the research area of the communication networks. NC has been firstly introduced by Ahlswede et al., who showed that NC can achieve the maximum-flow minimum-cut bound of the multicast [1]. Specifically, NC divides the original content into fixed size blocks and each node in the network can encode the received coded content blocks which belong to the same content.

With linear network coding (LNC), both the encoding and decoding operations are linear combinations on the received content blocks [2]. Li et al. proved that LNC is sufficient to achieve the maximum throughput of the multicast [2]. With low computational complexity, LNC has been widely studied in the literature because that it has advantages of increasing network throughput, improving transmission robustness, providing data confidentiality, etc. [1]–[4].

Nowadays, traditional multicast has been widely used in content delivery including video conference, video on demand, IP TV, etc. With LNC, the coded content blocks transmitted on each link are useful for different users which consequently can further improve the network throughput [2]. Moreover, with LNC, the transmission topology becomes a subgraph of the network topology (will be shown in Fig. 1) [2]–[7].

In addition to improving the network throughput, it has shown that LNC can naturally provide confidentiality of the content delivery [4], [8]–[11]. Specifically, when an attacker cannot acquire sufficient number of coded content blocks, it cannot decode any original content block [8]–[11]. To provide confidentiality, there are mainly two kinds of passive attacks, i.e., the outside wiretapping attack [4], [8] and the inside eavesdropping attack [10]–[12]. For the outside wiretapping attack, an attacker can wiretap a set of links and obtain all the data transmitted on these links. On the other hand, for the inside eavesdropping attack, a set of intermediate nodes attempt to gain as much information as possible from the content delivery passing through them. In this paper, we will consider the inside eavesdropping attack model.

To provide confidentiality, previous studies mainly focus on two types of secure requirements, i.e., information theoretical security [9], [13] and weak security [4], [10], [14], [15]. Information theoretical security does not allow any information related to the original content to be leaked to the attacker. On the other hand, weak security requires that the attackers cannot obtain any meaningful information of the original data. Since coded data blocks are transmitted on links of the network, LNC can be designed to make sure that the attackers cannot obtain any meaningful information of the original data without using traditional encryption/decryption [4], [10], [14]. In this paper, we will consider the secure LNC design to achieve the requirements of weak security for multicast.

In this paper, we consider a multicast with a source and a set of destinations, where all destinations receive content blocks of the same content at the same rate. Therefore, the rate of the

†Corresponding author: Jin Wang, wjin1985@suda.edu.cn.
This work was supported in part by Collaborative Innovation Center of Novel Software Technology and Industrialization, the National Natural Science Foundation of China (No. 61672370, 61672282 and 61373164), the China Postdoctoral Science Foundation (No.2013M531402 and 2014T70544), Natural Science Foundation of the Higher Education Institutions of Jiangsu Province (No. 16KJB520040) and the Application Foundation Research of Suzhou (No. SYG201401).

1We will formally define the concept of weak security in Sec. II-B.
multicast is limited by the minimum rate that the destination can receive content blocks. The secure transmission rate (STR) of a content is the maximum reachable multicast rate under the requirements of weak security. In this paper, we aim to maximize the STR for a multicast.

Since the design of LNC is based on the transmission topology, transmission topology construction has a great influence on the performance of LNC designs [3], [5]–[7]. Specifically, the transmission topology of a multicast includes a set of nodes and links involved in the multicast and the rate of data packets which will be transmitted on these links. [16] considered the transmission topology construction problem of a multicast to minimize total transmission cost (TC), which allows that the rate on each link is a real number. However, in fact, with LNC, the coded data packets transmitted on each link are fixed size. Therefore, it is more reasonable to consider that the rate under each link should be non-negative integer number, which denotes integral number of coded data packets transmitted on it. Under this consideration, [3], [5]–[7] studied the min-cost integral transmission topology construction problem for single multicast and multiple multiscasts. However, compared with existing works, the main differences of our work are (1) we consider that the link rates are constrained to be integers and (2) we jointly study the transmission topology construction and the LNC design together to maximize the STR and minimize the TC under the requirements of weak security.

Next, in Fig. 1, we give an example to illustrate the impacts of transmission topology construction and secure LNC design on the STR and the TC for the content delivery. Fig. 1 (a) shows a network topology with unit link transmission capacity and unit link transmission cost. The source s has content f, and destinations t1, t2 request the content. The original data blocks of content f are denoted as x1, x2, x3. We consider the inside eavesdropping attack where all the intermediate nodes attempt to gain as much information as possible from the coded data blocks passing through them individually [10]–[13], [15]. Moreover, LNC over finite field \( \mathbb{F}_2 \) is applied on data blocks and the coded data blocks are shown on the links.

- Fig. 1 (b) shows a transmission topology, in which each destination node can receive three data blocks in each transmission round. However, since all the data blocks transmitted to t1 and t2 must pass through the node v, it can decode all the original data blocks as t1 and t2 do. Therefore, there does not exist secure LNC on the transmission topology and the achieved STR is 0.
- In Fig. 1 (c), transmission topology 2 and the LNC scheme shows that each intermediate node cannot decode and obtain any original data blocks. Moreover, each destination can also decode and obtain three data blocks in each transmission round. In this case, the achieved STR is 3 and the TC of transmission topology 2 is 16. Therefore, comparing with Fig. 1 (b), we know transmission topology can determine whether it exists a secure LNC or not.
- As shown in Fig. 1 (d), each intermediate node cannot decode and obtain any original data blocks and each destination can also decode and obtain three data blocks in each transmission round. In this case, the STR is 3. However, the TC of transmission topology 3 is 12. Compared with Fig. 1 (c), it decreases 25%.

The above example shows that the transmission topology construction and the LNC design have significant impacts on the STR and the TC of content delivery. To the best of the authors’ knowledge, no previous works have been conducted to discuss the problem of optimal transmission topology construction and LNC design for the multicast with integral link rates under the requirements of weak security. The main contributions of the paper are summarized as follows:

- **Problem modeling**: We propose and model the problem of optimal transmission topology construction and secure LNC design for multicast with integral link rates, which is referred to as Integral Secure Multicast (ISM) problem.
- **Theoretical analysis**: We theoretically analyze the sufficient and necessary condition that there exists a transmission topology with integral link rates and a secure LNC.
- **Problem formulation and algorithm design**: Based on the sufficient and necessary condition, we formulate the ISM problem into an integer linear programming (ILP). We also design an efficient transmission topology construction (TTC) algorithm based on Lagrangian relaxation and subgradient algorithm to obtain a near-optimal solution.
- **Performance evaluation**: We conduct extensive experiments to evaluate the performance of the proposed TTC algorithm together with the lower bound and upper bound of the ISM problem. Simulation results validate the effectiveness of the proposed TTC algorithm.

The construction of the paper is organized as follows. We firstly introduce the system model and describe the problem specifically and systematically in the Sec. II. We then give theoretical analysis to show a sufficient and necessary condition that there exists a transmission topology with integral link rates and a secure LNC with STR \( R \) in Sec. III. Based on the condition, we then formulate the ISM problem into an
**Integral Secure Multicast**

In this subsection, to facilitate the definitions, we assume that each link has a unit capacity, which is 1 data unit per time slot. Note that we can always regard a link with capacity \( p \) data units per time as \( p \) links, each of which has a unit capacity.

**Definition 1.** \( \pi(N, s, T, F_q, R) \) is a \( R \)-dimensional LNC for multicast, which consists of GEVs with length \( R \), such that

- For the source \( s \), \( \bigcup_{w \in \mathcal{L}_s^+} \{\sigma_{s,w}\} \) contains a basis of the vector space \( F_q^R \).
- For each node \( v \in A - \{s\}, \sigma_{v,w} \) is a linear combination of vectors in \( \bigcup_{u \in \mathcal{L}_v^+} \{\sigma_{u,v}\}, \forall w \in \mathcal{L}_v^+ \).
- For each destination node \( t \in T \), \( \bigcup_{u \in \mathcal{L}_t^+} \{\sigma_{u,t}\} \) contains a basis of the vector space \( F_q^R \).

We note that a LNC over finite field \( F_q \) means that elements of all the GEVs and coded data blocks belong to \( F_q \). Moreover, the linear encoding and decoding are also processed within \( F_q \).

**B. Attack Model and Secure Requirements**

In this subsection, we introduce the attack model and secure requirements studied in this paper. For a multicast, we consider the eavesdropping attack where all the intermediate nodes in \( I \) attempt to gain as much information as possible from the coded data blocks passing through them individually and do not collude with each other [10]–[13], [15]. Specifically, when a coded data packet passes through a node, it obtains the GEV and the corresponding coded data block. Let the matrix composed by the GEVs obtained by a node \( w \in I \) as its rows be \( C_w \). The coded data blocks obtained by node \( w \) is \( C_w X \).

In this paper, we consider the weak security, which has been proposed in [14] and widely studied in [4], [10], [14], [15]. When transmitting a single content in the multicast and there is no correlations between the original data blocks, the weak security can be defined as follows:

**Definition 2.** A \( R \)-dimensional LNC \( \pi(N, s, T, F_q, R) \) for multicast is weakly secure, if it satisfies:

\[
H(x_j | C_w X) = H(x_j), \forall w \in I, \forall j \in \{1, \ldots, R\},
\]

where \( H(\cdot) \) and \( H(\cdot | \cdot) \) denote the entropy and conditional entropy, respectively [14].

Eq. (1) means that each intermediate node \( w \) cannot obtain any original data block \( x_j \) from the received coded data blocks \( C_w X, \forall j \in \{1, \ldots, R\} \). In other words, \( x_j \) does not belong to the space of row vectors of \( C_w X, \forall w \in I, \forall j \in \{1, \ldots, R\} \), i.e., the space of row vectors of \( C_w \) does not contain any row vector of the identity matrix with dimension \( R \times R \), \( \forall w \in I \).

In a network \( N \), if there exists a transmission topology and a \( R \)-dimensional LNC \( \pi \) which satisfies the requirements of weak security, we regard the transmission rate of the multicast as secure transmission rate (STR) and the STR is \( R \).

**Definition 3.** A \( R \)-dimensional LNC \( \pi(N, s, T, F_q, R) \) is a weakly Secure LNC for Multicast (SLM), if \( \pi \) is weakly secure.

**C. Problem Definition**

For a given multicast, in this paper, we study the problem of optimal transmission topology construction and secure LNC design when the link rates are constrained to be integers, which is referred to as the following Integral Secure Multicast (ISM) problem.
Definition 4. Given the network $N(A, L)$, a source node $s$ and a set of destinations $T$, the ISM problem aims to find the transmission topology with integral link rates and design the corresponding SLM to (1) meet the requirements of weak security (2) achieve the maximum STR and (3) minimize TC under the maximum STR guarantee.

III. THEORETICAL ANALYSIS

In this section, we will give theoretical analysis to show the sufficient and necessary condition that there exists a SLM which can achieve STR $R$. We firstly show the sufficient and necessary condition on the set of GEVs collected by each intermediate node.

Lemma 1. In $N$, if there exists a transmission topology with integral link rates and a SLM $\pi$ with STR $R$, then $\text{Rank}(C_w) \leq R - 1, \forall w \in I$.

Proof. Since the length of the each GEV is $R$, $\text{Rank}(C_w) \leq R$, $\forall w \in I$. Next, we prove the Lemma by contradiction. If there exists an intermediate node $w \in I, \text{Rank}(C_w) = R$, then $w$ can decode and obtain all the original data blocks which is in contradiction with the condition that $\pi$ is SLM. Therefore, we have $\text{Rank}(C_w) \leq R - 1, \forall w \in I$.

Lemma 2. In $N$, if there exists a $R$-dimensional LNC with $\text{Rank}(C_w) \leq R - 1, \forall w \in I$, there exists a transmission topology with integral link rates and a SLM $\pi$ with STR $R$.

Proof. The proof of existence of SLM $\pi$ with STR $R$ can be found in Theorem 1 of [14]. For the transmission topology, based on the SLM $\pi$, it contains all the links on which coded data blocks in $\pi$ are transmitted. Moreover, the rate of each link in the transmission topology is the number of coded data blocks transmitted on it per transmission round. Thus, the transmission topology is obviously with integral link rates.

Theorem 1. In $N$, there exists a transmission topology with integral link rates and a SLM $\pi$ with STR $R$, if and only if there exists a $R$-dimensional LNC with $\text{Rank}(C_w) \leq R - 1, \forall w \in I$.

Proof. From Lemma 1 and 2, it obviously holds.

According to [2], in $N$, there exists a LNC $\pi(N, s, T, F_q, R)$ for multicast from $s$ to $T$ if and only if there exists a network flow with flow capacity $R$ from $s$ to each destination. In LNC $\pi$, the number of coded data blocks transmitted on each link is no less than the actual flow rate, which is the maximum flow rate passing through the link. The reason is that with LNC, the coded data packets transmitted on each link can be useful and shared for all the destinations. Let $r_{u,v}$ be the actual flow rate on the link $l_{u,v}$. Next, we show the sufficient and necessary condition on the network flows that there exists a transmission topology with integral link rates and a SLM with STR $R$.

Lemma 3. In $N$, if there exists a network flow with flow capacity $R$ from $s$ to each destination and for each intermediate node $w \in I$, the total actual flow rate entering it is no more than $R - 1$, then there exists a transmission topology with integral link rates and a SLM $\pi$ with STR $R$.

Proof. If there exists a network flow with flow capacity $R$ from $s$ to each destination and for each intermediate node $w \in I$, the total actual flow rate entering it is no more than $R - 1$, according to [2], there exists a LNC for multicast from $s$ to $T$ in $N$. Moreover, since the number of coded data blocks transmitted on each link is no more than the actual flow of it and for each intermediate node $w \in I$, the total number of coded data packets received by it is no more than $R - 1$. We have $\text{Rank}(C_w) \leq R - 1, \forall w \in I$. Therefore, according to Theorem 1, there exists a transmission topology with integral link rates and a SLM $\pi$ with STR $R$.

Lemma 4. In $N$, if there exists a transmission topology with integral link rates and a SLM $\pi$ with STR $R$, then there exists a network flow with flow capacity $R$ from $s$ to each destination and for each intermediate node $w \in I$, the total actual flow rate entering it is no more than $R - 1$.

Proof. If there exists a SLM $\pi$ with STR $R$, then $\text{Rank}(C_w) \leq R - 1, \forall w \in I$ (Theorem 1). Let $C_w$ be the maximum linearly independent group of row vectors of $C_w$. For each intermediate node $w \in I$, if the number of coded data packets received is more than $R - 1$, we can remove the coded data packets with GEVs that are not in $C_w$ from the incoming links of $w$. Since all the outgoing data blocks can also be generated by the data blocks with GEVs in $C_w$, after removing such coded data packets, all the destinations can also decode and recover $R$ original data blocks. Moreover, such an LNC $\pi'$ is also weakly secure, because we only remove a set of coded data packets from $\pi$. In $\pi'$, the number of coded packets received by each intermediate node $w \in I$ is $|C_w|$ and $|C_w| = \text{Rank}(C_w) \leq R - 1$. According to [2], there exists a network flow with flow capacity $R$ from $s$ to each destination and for each intermediate node $w \in I$, the total actual flow rate entering it is no more than $R - 1$.

Theorem 2. In $N$, there exists a transmission topology with integral link rates and a SLM $\pi$ with STR $R$, if and only if there exists a network flow with flow capacity $R$ from $s$ to each destination and for each intermediate node $w \in I$, the total actual flow rate entering it is no more than $R - 1$.

Proof. From Lemma 4 and 5, it obviously holds.

Based on the above theoretical analysis, we present a design framework to optimally solve the ISM problem as follow:
1) Given $N$, $s$ and $T$, based on Theorem 2, we formulate the problem of finding the optimal transmission topology with integral link rates into an integer linear programming (ILP) $P$, which can be used to obtain the optimal solution, and also design a near-optimal transmission topology construction (TTC) algorithm to solve ILP $P$ based on Lagrangian relaxation and subgradient algorithm within polynomial time, which are shown in Sec. IV.

2) Based on the transmission topology of $N$, a LNC $\pi$ for the multicast with a source $s$ and a set of destinations $T$ can be designed according to [2] (shown in Fig. 2 (a)) and then a transformation matrix $D$ can be constructed (shown in Fig. 2 (b)) according to Theorem 1 in [14]. When the source $s$ transmits $DX$ instead of $X$, the LNC $\pi$ becomes a SLM $\pi'$ (shown in Fig. 2 (c)), which can achieve the same STR and TC as these acquired in Sec. IV.

IV. TRANSMISSION TOPOLOGY CONSTRUCTION ALGORITHM

Without security consideration, the problem of constructing transmission topology with integral link rates has been proved as an NP hard problem [3], [5]–[7]. In this section, we will firstly model the ISM problem to be an integer linear programming (ILP) $P$, which can be used to obtain the optimal solution when the size of the problem is small. If network parameters are large, based on Lagrangian relaxation and subgradient method, we design an efficient transmission topology construction (TTC) algorithm to solve the ILP $P$.

A. ILP Formulation for the ISM Problem

Based on the theoretical analysis shown in Sec. III, the ISM problem can be modeled as an ILP $P$ shown as follows:

$$\begin{align*}
\text{Maximize} & \quad R \\
\text{s.t.} & \quad \sum_{\{u|u \in A^v\}} r^t_{u,v} - \sum_{\{u|u \in L^v\}} r^t_{v,u} = R_t, v = t, \forall t \in T \quad (3) \\
& \quad \sum_{\{u|u \in A^v\}} r^t_{u,v} - \sum_{\{u|u \in L^v\}} r^t_{v,u} = -R_t, v = s, \forall t \in T \quad (4) \\
& \quad \sum_{\{u|u \in A^v\}} r^t_{u,v} - \sum_{\{u|u \in L^v\}} r^t_{v,u} = 0, \forall v \in A - \{s,t\}, \forall t \in T \quad (5) \\
& \quad r^t_{u,v} \leq R_{uv}, \forall u, v \in L, \forall t \in T \quad (6) \\
& \quad r^t_{u,v} \leq R - 1, \forall v \in I, I = A - \{s\} - T \quad (7) \\
& \quad 0 \leq r^t_{u,v} \leq p_{u,v}, \forall u, v \in L \quad (8) \\
& \quad 0 \leq r^t_{u,v} \leq \rho_{u,v}, \forall u, v \in L, \forall t \in T \quad (9) \\
& \quad R, r_{u,v}, r_{u,v} \in Z, \forall u, v \in L, \forall t \in T \quad (10) \\
& \quad \sum_{\{u|u \in L^v\}} r^t_{v,u} \leq r_{uv}, \forall u, v \in L, \forall t \in T \quad (11) \\
\end{align*}$$

The objective (2) is to maximize the STR $R$. $r^t_{u,v}$ denotes the rate of flow from the source node $s$ to the destination node $t$ on the link $l_{u,v}$. Constraints (3)–(5) show the network flow conservation constraints on the flows from $s$ to all the destinations. Specifically, $R_t$ denotes the flow capacity of each flow from $s$ to each destination $t$. Therefore, the receiving rate of destination $t$ is $R_t$. $r_{u,v}$ denotes the actual flow rate transmitted on link $l_{u,v}$. With LNC, the coded data packets transmitted on each link can be shared for all the destinations. Therefore, the number of coded data packets transmitted on link $l_{u,v}$ only needs to be no less than $\max_{t \in T} r^t_{u,v}$, which is shown in constraint (6). According to Theorem 2, constraint (7) shows the security constraint on each intermediate node. Constraint (8) means that $R \leq \min_{t \in T} R_t$. Considering the objective (2), the constraint (8) equals to $R = \min_{t \in T} R_t$, which is the STR of the multicast. Constraint (9) shows the link capacity constraint. Constraints (10)–(11) enforce the value ranges of $r^t_{u,v}$ and $r_{u,v}$. Specifically, the actual flow rate $r_{u,v}$ on each link $l_{u,v}$ should be integers.

After we obtain the maximum STR $R^*$, by setting STR $R^*$ as a known parameter, we can further solve the following ILP $P'$ to obtain the transmission topology with minimum TC. Then, a SLM can be designed to achieve STR $R^*$.

$$\begin{align*}
\text{Minimize} & \quad \sum_{l_{u,v} \in L} c_{u,v} r_{u,v} \\
\text{s.t.} & \quad \sum_{\{u|u \in L^v\}} r^t_{v,u} - \sum_{\{u|u \in A^v\}} r^t_{u,v} = R_t, v = t, \forall t \in T \quad (13) \\
& \quad R_t = R^*, \forall t \in T \quad (14)
\end{align*}$$

The proposed ILP $P$ and $P'$ can be used to directly obtain the optimal transmission topology of the ISM problem when the size of the problem is small.

B. An Efficient Transmission Topology Construction (TTC) Algorithm

In this section, based on Lagrangian relaxation and subgradient algorithm, we design an efficient transmission topology construction (TTC) algorithm to obtain a near-optimal solution of the ISM problem.

1) ILP $P_1$ for a given STR $R$:

Before we give the TTC algorithm, we first consider the problem of finding transmission topology with minimum TC. Then, based on the topology, a SLM can be designed to achieve a given STR $R$. In this case, the problem can be formulated into an ILP as follows:

$$\begin{align*}
\text{Minimize} & \quad \sum_{l_{u,v} \in L} c_{u,v} r_{u,v} \\
\text{s.t.} & \quad \sum_{\{u|u \in L^v\}} r^t_{v,u} - \sum_{\{u|u \in A^v\}} r^t_{u,v} = R_t, v = t, \forall t \in T \\
& \quad -R \leq v = s, \forall t \in T \\
& \quad 0 \leq v \leq A - \{s,t\}, \forall t \in T \\
& \quad r^t_{u,v} \leq r_{uv}, \forall u, v \in L, \forall t \in T \\
& \quad 0 \leq r^t_{u,v} \leq p_{u,v}, \forall u, v \in L \quad (15) \\
& \quad 0 \leq r^t_{u,v} \leq \rho_{u,v}, \forall u, v \in L, \forall t \in T \\
& \quad \sum_{\{u|u \in L^v\}} r^t_{v,u} \leq R - 1, \forall v \in I, I = A - \{s\} - T \\
& \quad 0 \leq r^t_{u,v} \leq p_{u,v}, \forall u, v \in L, \forall t \in T \\
& \quad 0 \leq r^t_{u,v} \leq \rho_{u,v}, \forall u, v \in L, \forall t \in T \\
& \quad r_{u,v} \leq \rho_{u,v}, \forall u, v \in L, \forall t \in T \\
& \quad r_{u,v} \leq p_{u,v}, \forall u, v \in L, \forall t \in T \\
& \quad r^t_{u,v} \leq Z, \forall u, v \in L, \forall t \in T \\
& \quad r_{u,v} \in Z, \forall u, v \in L \\
\end{align*}$$
2) Lagrangian Dual Problem of ILP $P_1$: Based on Lagrangian relaxation and subgradient algorithm, we then design an efficient algorithm to solve ILP $P_1$. Specifically, we first move the constraint (17) to the objective function and the $p_{u,v,t}, \forall u,v \in L, \forall t \in T$ are Lagrangian multipliers. The objective function of ILP $P_1$ becomes:

$$\sum_{i,u,v \in L} c_{u,v} r_{u,v} + \sum_{i,u,v \in L, t \in T} \rho_{u,v,t} (r_{u,v}^t - r_{u,v}).$$

Let $\rho$ be the vector which consists of $\rho_{u,v,t}, \forall u,v \in L, \forall t \in T$. The Lagrangian dual problem of ILP $P_1$ is shown as follows:

$$\max_{\rho>0} \Gamma(\rho),$$

in which

$$P_\rho: \Gamma(\rho) = \min \sum_{i,u,v \in L} c_{u,v} r_{u,v} + \sum_{i,u,v \in L, t \in T} \rho_{u,v,t} (r_{u,v}^t - r_{u,v})$$

s.t.

Constraints (16), (18) – (22)

Given $\rho$, the result of above Lagrangian relaxation of ILP $P_\rho$ is a lower bound of the original ILP $P_1$ [17]. Next, we prove the coefficient matrix of the constraints of $P_\rho$ is a totally unimodular matrix [18]. Let $M$ denotes the coefficient matrix of the constraints (16) and (18).

**Theorem 3.** $M$ is a totally unimodular matrix.

**Proof.** $M$ is a totally unimodular matrix, when it meets the following three conditions [18].

For the first condition that every element of the matrix should be $0$, $-1$ or $+1$. The matrix $M$ obviously satisfies.

The second condition is that each column of the matrix should have no more than two non-zero elements. For the matrix $M$, it means that each variable appears at most two times in the constraints (16) and (18). There are two sets of variables: $\{r_{u,v}^t\}$ and $\{r_{u,v}\}$. For given $t \in T, v \in A$ and $u \in \lambda^t_v$, variable $r_{u,v}$ appears two times only in the constraint (16), because the link $l_{u,v}$ can be an incoming link for node $v$ and an outgoing link for node $u$. Specifically, the coefficients of it are $+1$ for the incoming link for node $v$ and $-1$ for the outgoing link for node $u$, respectively. On the other hand, for given $v \in I$ and $u \in \lambda_v^t$, variable $r_{u,v}$ appears one time only in the constraint (18) because a link $l_{u,v}$ can only be an incoming link for only one node $v$. Specifically, the coefficients of the variable $r_{u,v}$ is $+1$ when it appears in constraints (18).

The third condition is that if the two non-zero elements in a column of $M$ have the same sign, then one of the rows is in set $S_1$, and the other in set $S_2$, and if the two non-zero elements in a column of $M$ have opposite signs, then the rows of both are in set $S_1$, or both in set $S_2$. As above discussion, firstly, since for each variable $r_{u,v}^t$, the corresponding column must contains one $+1$ and one $-1$. We can set all the rows of $M$, which are related to the constraint (16), belong to set $S_1$. Then, for each variable $r_{u,v}$, the corresponding column only contains one $+1$. We can set all the rows of $M$, which are related to the constraint (18), belong to set $S_2$. In this case, $M$ satisfies the third condition.

Although $P_\rho$ is also an ILP, it has integral optimal solution when we remove the integer constraints (21) and (22), because the coefficient matrix of $P_\rho$ has been proved to be a totally unimodular matrix (Theorem 19.3 in [19]). Therefore, when giving $\rho$, $P_\rho$ can be solved within polynomial computational complexity, i.e., $P_\rho$ is equivalent to the following LP $P_\rho^L$:

$$P_\rho^L: \min \sum_{i,u,v \in L} c_{u,v} r_{u,v} + \sum_{i,u,v \in L, t \in T} \rho_{u,v,t} (r_{u,v}^t - r_{u,v})$$

s.t.

Constraints (16), (18) – (22)

3) Selection of Multipliers for the Lagrangian Dual Problem: To solve the Lagrangian dual problem shown in (23), it is important to select an appropriate multiplier vector $\rho$. The subgradient optimization is commonly used in $\rho$ selection [17].

The multiplier vector $\rho$ of the $(i+1)^{th}$ iteration is denoted as $\rho^{i+1}$ and $\rho^{i+1} = \max(\rho^i + \delta \epsilon, 0)$, in which $\delta_1$ is a positive step size. Let $\delta_1 = \theta_1(iU_B - 1 T_L)$, in which $0 \leq \theta_i \leq 2$. At iteration $i$, the subgradient vector $\epsilon$ consists of elements $\epsilon_{u,v,t} = r_{u,v}^t - r_{u,v}$, and the notations $r_{u,v}$ denote the values of variables $r_{u,v}$, and $r_{u,v}^t$ for the outgoing link for node $v$ and the incoming link for node $u$, respectively. On the other hand, for given $v \in I$ and $u \in \lambda_v^t$, variable $r_{u,v}$ appears one time only in the constraint (18) because a link $l_{u,v}$ can only be an incoming link for only one node $v$. Specifically, the coefficients of the variable $r_{u,v}$ is $+1$ when it appears in constraints (18).

The third condition is that if the two non-zero elements in a column of $M$ have the same sign, then one of the rows is in set $S_1$, and the other in set $S_2$, and if the two non-zero elements in a column of $M$ have opposite signs, then the rows of both are in set $S_1$, or both in set $S_2$. As above discussion, firstly, since for each variable $r_{u,v}^t$, the corresponding column must contains one $+1$ and one $-1$. We can set all the rows of $M$, which are related to the constraint (16), belong to set $S_1$. Then, for each variable $r_{u,v}$, the corresponding column only contains one $+1$. We can set all the rows of $M$, which are related to the constraint (18), belong to set $S_2$. In this case, $M$ satisfies the third condition.
Applying Ford-Fulkerson method from the source to each link cost is randomly selected from.

The designs of TTC algorithm are shown in Alg. 1 and Alg. 2. Specifically, Alg. 1 gives no solution if (1) for all $\Gamma^{1}_{UB}$ and $\Gamma^{1}_{LB}$ has no feasible solution then return $I = 0, NULL$; or (2) $\Gamma^{1}_{UB} > \Gamma^{1}_{LB}$ then return $\Gamma^{1}_{UB} = 1$; else obtain $\Gamma^{1}_{UB} = \Gamma^{1}_{LB}$; or (3) $\Gamma^{1}_{UB} > \Gamma^{1}_{LB}$ then return $\Gamma^{1}_{UB} = \Gamma^{1}_{LB}$.

4) Transmission Topology Construction (TTC) Algorithm:
The designs of TTC algorithm are shown in Alg. 1 and Alg. 2. Specifically, in Alg. 1, we firstly obtain the maximum transmission rate $r$ from the source $s$ to all destinations by applying Ford-Fulkerson method from the source to each destination. We know that the maximum STR of the ISM problem is no more than $r$. Therefore, we verify that when the value of $R$ is given from $r$ to 1, whether there exists a transmission topology to achieve the STR $R$, which is processed in Alg. 2. Alg. 2 realizes the Lagrangian relaxation and subgradient algorithm which solves problem $P_1$ when STR $R$ is given. The Alg. 1 give no solution, if (1) for all values of $R$, Alg. 2 gives no solution; or (2) $R = 1$. When $r = 1, R \leq 1$, which means no transmission topology and SLM exists according to Theorem 2. Specifically, Alg. 2 will stop if: (1) the number of iteration $i \geq T$; or (2) the difference between $\Gamma^{1}_{UB}$ and $\Gamma^{1}_{LB}$ is less than a threshold $\alpha^*$.

V. SIMULATION

In this section, we evaluate the performance of the proposed TTC algorithm in random networks generated by a widely used Waxman model [20]. For the STR of a multicast, the maximum transmission rate without considering security requirements is obviously an upper bound, which is denoted as UB. Specifically, UB can be obtained from ILP $P$ without the constraints (6), (7), (9) and (11). From Sec. IV, a lower bound of the minimum TC obtained by the TTC algorithm can be obtained by relaxing the integer constraints (21) and (22) of ILP $P_1$. We denote it as LB.

The region of the random network $N = 10 \times 10$. The network parameters of the Waxman network model are introduced as follows: the intensity of Poisson process $\lambda = 0.4$, the maximal link probability $\alpha = 0.4$, and the parameter to control length of the links $\beta = 0.4$.

There are three parameters in the following simulations.

- $|T|$: the number of the destination nodes, $|T| \in [1, 10]$,
- $Ca$: the maximum link capacity, $Ca \in [1, 10]$,
- $Co$: the maximum link cost, $Co \in [1, 10]$.

For each link, the capacity is randomly selected from $[1, Ca]$ and link cost is randomly selected from $[1, Co]$. For a network without the STR of a multicast, the network parameters of the Waxman network model are introduced as follows: the intensity of Poisson process $\rho = 10^{-5}$, $\forall u,v \in L, \forall t \in T$ and $o' = +\infty$.

For each link, the capacity is randomly selected from $[1, Ca]$ and link cost is randomly selected from $[1, Co]$. For a network without the STR of a multicast, the network parameters of the Waxman network model are introduced as follows: the intensity of Poisson process $\rho = 10^{-5}$, $\forall u,v \in L, \forall t \in T$ and $o' = +\infty$.

V. SIMULATION

In this section, we evaluate the performance of the proposed TTC algorithm in random networks generated by a widely used Waxman model [20]. For the STR of a multicast, the maximum transmission rate without considering security requirements is obviously an upper bound, which is denoted as UB. Specifically, UB can be obtained from ILP $P$ without the constraints (6), (7), (9) and (11). From Sec. IV, a lower bound of the minimum TC obtained by the TTC algorithm can be obtained by relaxing the integer constraints (21) and (22) of ILP $P_1$. We denote it as LB.

The region of the random network $N = 10 \times 10$. The network parameters of the Waxman network model are introduced as follows: the intensity of Poisson process $\lambda = 0.4$, the maximal link probability $\alpha = 0.4$, and the parameter to control length of the links $\beta = 0.4$.

There are three parameters in the following simulations.

- $|T|$: the number of the destination nodes, $|T| \in [1, 10]$,
- $Ca$: the maximum link capacity, $Ca \in [1, 10]$,
- $Co$: the maximum link cost, $Co \in [1, 10]$.

For each link, the capacity is randomly selected from $[1, Ca]$ and link cost is randomly selected from $[1, Co]$. For a network without the STR of a multicast, the network parameters of the Waxman network model are introduced as follows: the intensity of Poisson process $\rho = 10^{-5}$, $\forall u,v \in L, \forall t \in T$ and $o' = +\infty$.

V. SIMULATION

In this section, we evaluate the performance of the proposed TTC algorithm in random networks generated by a widely used Waxman model [20]. For the STR of a multicast, the maximum transmission rate without considering security requirements is obviously an upper bound, which is denoted as UB. Specifically, UB can be obtained from ILP $P$ without the constraints (6), (7), (9) and (11). From Sec. IV, a lower bound of the minimum TC obtained by the TTC algorithm can be obtained by relaxing the integer constraints (21) and (22) of ILP $P_1$. We denote it as LB.

The region of the random network $N = 10 \times 10$. The network parameters of the Waxman network model are introduced as follows: the intensity of Poisson process $\lambda = 0.4$, the maximal link probability $\alpha = 0.4$, and the parameter to control length of the links $\beta = 0.4$.

There are three parameters in the following simulations.

- $|T|$: the number of the destination nodes, $|T| \in [1, 10]$,
- $Ca$: the maximum link capacity, $Ca \in [1, 10]$,
- $Co$: the maximum link cost, $Co \in [1, 10]$.

For each link, the capacity is randomly selected from $[1, Ca]$ and link cost is randomly selected from $[1, Co]$. For a network without the STR of a multicast, the network parameters of the Waxman network model are introduced as follows: the intensity of Poisson process $\rho = 10^{-5}$, $\forall u,v \in L, \forall t \in T$ and $o' = +\infty$.
and the parameters of the link cost $Co$ are fixed. When $Ca$ increases from 1 to 10, Fig. 3 (c) shows both the STR of TTC and the transmission rate of UB increase. Fig. 3 (d) shows the same trend. The reason is that larger link capacity leads to higher STR and TC obviously. We note that the relative difference of the STR between UB and TTC is less than 4% and the relative difference of TC between TTC and LB is less than 7% even if $Co$ becomes sufficiently large ($Co \geq 10$).

In Fig. 3 (e), $|T|$ and $Ca$ are fixed. In this case, we do not show the STR of TTC and UB because they will not change obviously. However, the TCs of TTC and LB obviously increase when $Co$ increases. The relative difference of TC between TTC and LB is less than 9% even if $Co$ becomes sufficiently large ($Co \geq 10$).

Conclude from the simulations, the STR obtained by the proposed TTC algorithm is close to the UB and the relative difference between them is less than 4% in all cases. Besides, when the STR is maximized, the minimum TC increases with more destination nodes, larger link capacity or larger unit link cost. The relative difference between TCs achieved by TTC and LB is less than 12% even if $|T|, Ca, Co$ become sufficiently large.

VI. CONCLUSION

In this paper, we have investigated the Integral Secure Multicast (ISM) problem, which aims to find the transmission topology with integral link rates and design the SLM to (1) meet the requirements of weak security, (2) achieve the maximum STR and (3) minimize TC under the maximum STR guarantee. Specifically, in this paper, we firstly gave theoretical analysis to show a sufficient and necessary condition that there exists a transmission topology with integral link rates and a SLM $\pi$ with STR $R$. Based on the condition, we then formulated the ISM problem into an integer linear programming (ILP) and designed a near-optimal transmission topology construction (TTC) algorithm to solve the ILP within polynomial time, which is based on Lagrangian relaxation and subgradient algorithm. Finally, we simulated the performance of the proposed TTC algorithm together with the lower bound as well as upper bound of the ISM problem. Simulation results validated the effectiveness of the proposed TTC algorithm.

REFERENCES


