On Secure Network Coding with Unequal Link Capacities and Restricted Wiretapping Sets

Tao Cui and Tracey Ho
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125, USA
Email: {taocui, tho}@caltech.edu

Jörg Kliewer
Klipsch School of Electrical and Computer Engineering
New Mexico State University
Las Cruces, NM 88003, USA
Email: jkliewer@nmsu.edu

***Abstract*—We address secure network coding over networks with unequal link capacities in the presence of a wiretapper who has only access to a restricted number of k links in the network. Previous results show that for the case of equal link capacities and unrestricted wiretapping sets, the secrecy capacity is given by the cut-set bound, whether or not the location of the wiretapped links is known. The cut-set bound can be achieved by injecting k random keys at the source which are decoded at the sink along with the message. In contrast, for the case where the wiretapping set is restricted, or where link capacities are not equal, we show that the cut-set bound is not achievable in general. Finally, it is shown that determining the secrecy capacity is a NP-hard problem.**

I. Introduction

Information-theoretically secure communication uses coding to ensure that an adversary eavesdropping on a subset of network links obtains no information about the secure message. A theoretical basis for information-theoretic security was given in the seminal paper by Wyner [1] using Shannon's notion of perfect secrecy [2], where a coset coding scheme based on a linear maximum distance separable code was used to achieve security for a wiretap channel. More recently, information-theoretic security has been studied in networks with general topologies. The secure network coding problem was introduced in [3] for multicast wireline networks where each link has equal capacity, and a wiretapper can observe an unknown set of up to k network links. For this problem, constructions of information-theoretically secure linear network codes are proposed in e.g. [3]–[5], where trade-offs between security, code alphabet size, and multicast rate of secure linear network codes are considered in [4]. In [6], [7] the work in [3] is extended to multiple sources, where random keys can now be generated at an arbitrarily given subset of nodes. Further, in [8], secure communication is considered for wireless erasure networks.

In this paper, we consider secure communication over wireline networks with unequal link capacities and restricted wiretapping sets. In the case of throughput optimization without security requirements, the assumption that all links have unit capacity is made without loss of generality, since links of larger capacity can be modeled as multiple unit capacity links in parallel. However, in the secure communication problem, such an assumption cannot be made without loss of generality. Indeed, we show in this paper that there are significant differences between the equal capacity and unequal capacity cases. For the case of equal link capacities, the secrecy capacity is given by the cut-set bound, whether or not the location of the k wiretapped links is known. The cut-set bound can be achieved by injecting k random keys at the source which are decoded at the sink along with the message [3]. However, we show that if the network has unequal link capacities and the wiretapper has access to an unrestricted set of links in the network the cut-set bound is not achievable in general by any linear or nonlinear coding scheme. We further show that this also holds in the case of an unrestricted wiretapping set and equal unit link capacities in the network. Finally, we address the complexity of determining the secrecy capacity if the location of the wiretapper is unknown. We show that this problem, which is closely related to network interdiction, is NP-hard.

II. Network Model and Problem Formulation

In this paper we focus on acyclic graphs for simplicity; we expect that our results can be generalized to cyclic networks using the approach in [9], [10] of working over fields of rational functions in an indeterminate delay variable.

We model a wireline network by a directed acyclic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the vertex set and $\mathcal{E}$ is the directed edge set. There is a source node $s \in \mathcal{V}$ and a sink node $d \in \mathcal{V}$. Each link $(i, j) \in \mathcal{E}$ has capacity $c_{i,j}$.

An eavesdropper can wiretap a set $\mathcal{A}$ of links chosen from a known collection $\mathcal{W}$ of possible wiretap sets. Without loss of generality we can restrict our attention to maximal wiretap sets, i.e. no set in $\mathcal{W}$ is a subset of another. The choice of wiretap set $\mathcal{A}$ is unknown to the communicating nodes, except where otherwise specified in this paper. The secrecy capacity is the highest possible source-sink communication rate such that the message communicated is information theoretically secret regardless of the choice of $\mathcal{A}$, i.e. has zero mutual information with the wiretapper's observations.

In Sections III and IV, we show that the cut-set bound is unachievable and that finding the secrecy capacity is NP hard,

This work has been supported in part by subcontract #069144 issued by BAE Systems National Security Solutions, Inc. and supported by the Defense Advanced Research Projects Agency (DARPA) and the Space and Naval Warfare System Center (SPAWARSYSCEN), San Diego under Contracts No. N66001-08-C-2013 and W911NF-07-1-0029, by NSF grants CNS 0905615 and CCF 0830666, and by Caltech's Lee Center for Advanced Networking.

even for the following special cases:

1) Scenario 1 is a wireline network with *equal* link capacities, where the wiretapper can wiretap an unknown subset of k links from a known collection of vulnerable network links.
2) Scenario 2 is a wireline network with *unequal* link capacities, where the wiretapper can wiretap an unknown subset of k links from the entire network.

It is convenient to show these results for Scenario 1 first, and then show the corresponding results for Scenario 2, by converting the Scenario 1 networks considered into corresponding Scenario 2 networks for which the same result holds.

III. Unachievability of cut-set Bound

Let $\mathcal{S}^c$ denote the set complement of a set $\mathcal{S}$. A cut for $x, y \in \mathcal{V}$ is a partition of $\mathcal{V}$ into two sets $\mathcal{V}_x$ and $\mathcal{V}_x^c$ such that $x \in \mathcal{V}_x$ and $y \in \mathcal{V}_x^c$. For the $x-y$ cut given by $\mathcal{V}_x$, the cut-set $[\mathcal{V}_x, \mathcal{V}_x^c]$ is the set of edges going from $\mathcal{V}_x$ to $\mathcal{V}_x^c$, i.e.,

$$[\mathcal{V}_x, \mathcal{V}_x^c] = \{(u,v) | (u,v) \in \mathcal{E},\, u \in \mathcal{V}_x,\, v \in \mathcal{V}_x^c\}\,. \qquad (1)$$

We can state the following cut-set upper bound which applies to the general model of Section II including both scenarios 1 and 2:

Theorem 1. *Consider a network of point-to-point links, where link (i,j) has capacity $c_{i,j}$. The secrecy capacity R_s is upper bounded by*

$$\min_{\{\mathcal{V}_s:\, \mathcal{V}_s \text{ is an } s-d \text{ cut}\}} \; \min_{\mathcal{A}\in\mathcal{W}} \sum_{(i,j)\in[\mathcal{V}_s,\mathcal{V}_s^c]\cap\mathcal{A}^c} c_{i,j}\,. \qquad (2)$$

This upper bound applies whether or not the communicating nodes have knowledge of the chosen wiretap set $\mathcal{A}$.

Proof: Consider any source-sink cut $\mathcal{V}_s$ and any wiretap set $\mathcal{A} \in \mathcal{W}$. Denote by $\mathbf{X}$ the transmitted signals from nodes in $\mathcal{V}_s$ over links in $[\mathcal{V}_s, \mathcal{V}_s^c]$ and denote by $\mathbf{Y}$ and $\mathbf{Z}$ the observed signals from links in $[\mathcal{V}_s, \mathcal{V}_s^c]$ and in $[\mathcal{V}_s, \mathcal{V}_s^c] \cap \mathcal{A}$, respectively. We consider block coding with block length n. By the perfect secrecy requirement $H(M|\mathbf{Z}^n) = H(M)$ we have

$$\begin{aligned}
nR_s \leq & H(M|\mathbf{Z}^n)\\
\overset{(a)}{\leq} & H(M|\mathbf{Z}^n) - H(M|\mathbf{Y}^n) + n\epsilon_n\\
= & H(M|\mathbf{Z}^n) - H(M|\mathbf{Y}^n, \mathbf{Z}^n) + n\epsilon_n\\
= & I(M; \mathbf{Y}^n|\mathbf{Z}^n) + n\epsilon_n\\
\overset{(b)}{\leq} & I(\mathbf{X}^n; \mathbf{Y}^n|\mathbf{Z}^n) + n\epsilon_n\\
\overset{(c)}{\leq} & \sum_{i=1}^{n} H(Y_i|Z_i) - \sum_{i=1}^{n} H(Y_i|X_i, Z_i) + n\epsilon_n, \qquad (3)\\
= & nI(\mathbf{X}; \mathbf{Y}|\mathbf{Z}) + n\epsilon_n,\\
= & n\left(H(\mathbf{X}|\mathbf{Z}) - H(\mathbf{X}|\mathbf{Z}, \mathbf{Y})\right) + n\epsilon_n,\\
= & n\left(H(\mathbf{X}|\mathbf{Z}) - H(\mathbf{X}|\mathbf{Y})\right) + n\epsilon_n,\\
\leq & n \max_{p(\mathbf{X})} \left(I(\mathbf{X}; \mathbf{Y}) - I(\mathbf{X}; \mathbf{Z})\right) + n\epsilon_n\\
= & n \sum_{(i,j)\in[\mathcal{V}_s,\mathcal{V}_s^c]\cap\mathcal{A}^c} c_{i,j} + n\epsilon_n,
\end{aligned}$$

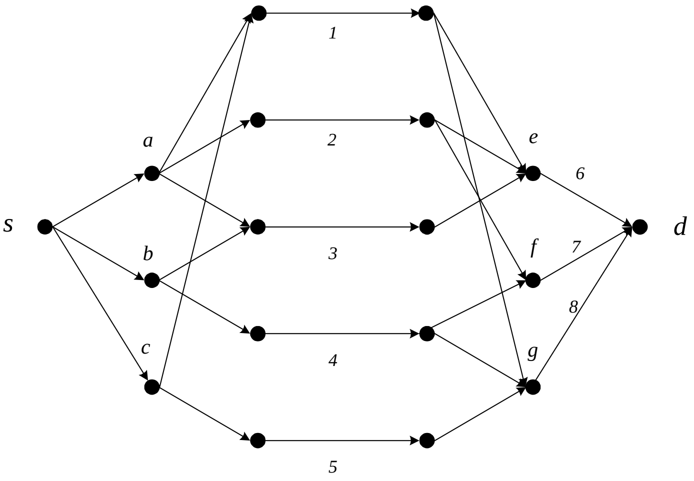


Fig. 1. An example to show that the secrecy rate without knowledge of wiretapping set is smaller than that with such knowledge. The wiretapper can wiretap any three of the five links in the middle layer.

where $\epsilon_n \to 0$ as $n \to +\infty$ and

- (a) is due to Fano's inequality;
- (b) is due to the data processing inequality and the fact that $M \to \mathbf{X}^n \to \mathbf{Y}^n \to \mathbf{Z}^n$ forms a Markov chain;
- (c) is due to the definition of the mutual information.

■

Note that Theorem 1 is a generalization of the bound in [6] to arbitrary link capacities.

If the choice of wiretap set $\mathcal{A}$ is known to the communicating nodes, the cut-set bound (2) is achievable using a network code that does not send any flow on links in $\mathcal{A}$. In the case of unrestricted wiretapping sets and unit link capacities, the secrecy capacity is equal to the cut-set bound [3]. In contrast, we now show that the cut-set bound is not achievable in general when the wiretap set $\mathcal{A}$ is unknown, by considering the example in Fig. 1, where the set of wiretappable links is restricted (Scenario 1). We use the program Information Theoretic Inequalities Prover (Xitip) [11] to show that the secrecy capacity is bounded away from the cut-set bound. We then convert the example into one with unequal link capacities (Scenario 2), and show the unachievability of the cut-set bound for this case also.

A. Restricted Wiretap Set (Scenario 1)

In Fig. 1, let the middle layer links be 1-5 (from top to bottom) and the last layer links be 6-8 (from top to bottom). All links have unit capacity. Let the signal carried by link i be called signal i, or S_i. Let the source information be denoted X. For this example, the secrecy rate is two if any three of the five links in the middle layer are deleted, i.e., the number of wiretapped links is three.

The constraints required are that the source information is a function of the signals on the sink's incoming links, and that there is zero mutual information between the source information and the signals on the links in each adversarial subset.

In this example, the cut-set bound is 2. To provide intuition, we first show that secrecy rate 2 cannot be achieved by using linear coding. This argument can be converted to an information theoretic proof that secrecy rate 2 cannot be achieved using any coding schemes [12].

Suppose secrecy rate 2 is achievable with a linear network code. First note that the source cannot inject more than unit amount of random key, otherwise the first layer cannot carry two units of source data. Let the random key injected by the source be denoted K. For the case when the source injects a unit amount of secret key, we first have the following observations. Signal 6 must be a function of signal 1, otherwise if the adversary sees the signals 2-4 then he knows signals 6-7. Also, signal 8 must be a function of signal 5, otherwise if the adversary sees signals 1, 2 and 4, then he knows signals 7-8. Similarly we can show that signal 8 must be a function of signal 1, and signal 7 must be a function of signal 2. We consider the following two cases.

Case 1: signal 5 is a linear combination of signals present at the source node. To achieve the full key rank condition on links 1, 2 and 5, the top second layer node (a) must put independent local keys k_1 and k_2 on links 1 and 2 respectively. Link 7, whose other input is independent of k_2, is then a function of k_2. Similarly, Link 8 is a function of k_1. This means that the last layer has two independent local keys on it.

Case 2: signal 5 is a linear combination of signals present at the source node as well as a local key k injected by the bottom second layer node (c).

Case 2a: k is also present in signal 1. Then k is present in signal 6, and is independent of the key present in signal 7.

Case 2b: k is not present in signal 1. Then k is present in signal 8, and is independent of the key present in signal 7.

From Cases 1, 2a, and 2b, we conclude that the secrecy rate without knowledge of the wiretapping set by using only linear network coding is less than two.

We can also show that the secrecy rate is bounded away from 2 by using the framework for linear information inequalities [13]. Let X be the message sent from the source and Z_i, $i = 1, \dots, 3$ be the signals on the links adjacent to the source. We want to check whether $H(X) \leq \omega$ is implied by

$$\begin{aligned}
&(1) && H(Z_i) \leq 1,\ H(S_j) \leq 1,\ i = 1, \dots, 3,\ j = 1, \dots, 8, \\
&(2) && H(X|S_6, S_7, S_8) = 0, \\
&(3) && I(X, Z_1, Z_2, Z_3, S_4, S_5, S_7, S_8; S_6|S_1, S_2, S_3) = 0, \\
&(4) && I(X, Z_1, Z_2, Z_3, S_1, S_3, S_5, S_6, S_8; S_7|S_2, S_4) = 0, \\
&(5) && I(X, Z_1, Z_2, Z_3, S_2, S_3, S_6, S_7; S_8|S_1, S_4, S_5) = 0, \\
&(6) && I(X; S_1, S_2, S_3) = 0,\ I(X; S_1, S_2, S_4) = 0, \\
&(7) && I(X; S_1, S_2, S_5) = 0,\ I(X; S_1, S_3, S_4) = 0, \\
&(8) && I(X; S_1, S_3, S_5) = 0,\ I(X; S_1, S_4, S_5) = 0, \\
&(9) && I(X; S_2, S_3, S_4) = 0,\ I(X; S_2, S_3, S_5) = 0, \\
&(10) && I(X; S_2, S_4, S_5) = 0,\ I(X; S_3, S_4, S_5) = 0, \\
&(11) && I(S_1; Z_2|Z_1, Z_3) = 0,\ I(S_2; Z_2, Z_3|Z_1) = 0, \\
&(12) && I(S_3; Z_3|Z_1, Z_2) = 0,\ I(S_4; Z_1, Z_3|Z_2) = 0, \\
&(13) && I(S_5; Z_1, Z_2|Z_3) = 0,\ I(S_1; S_4|Z_1, Z_2, Z_3) = 0, \\
&(14) && I(S_2; S_4, S_5|Z_1, Z_2, Z_3) = 0, \\
& && I(S_3; S_5|Z_1, Z_2, Z_3) = 0, \\
&(15) && I(S_4; S_1, S_2, S_5|Z_1, Z_2, Z_3) = 0, \\
& && I(S_5; S_2, S_3, S_4|Z_1, Z_2, Z_3) = 0, \\
&(16) && I(S_1, S_2, S_3, S_4, S_5; X|Z_1, Z_2, Z_3) = 0,
\end{aligned} \tag{4}$$

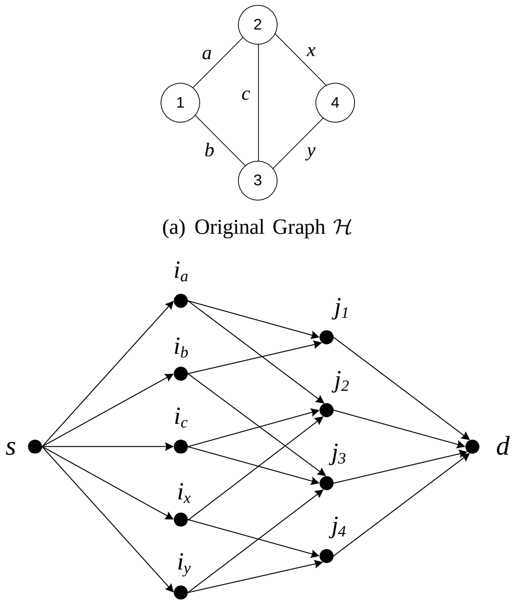
where the first inequality is the capacity constraint, the second constraint shows that the sink can decode X, constraints (3) to (5) mean that the signals in the last layer are independent of other signals given the incoming signals from the middle layer, constraints (6) to (10) represent the secrecy constraints when any three links in the middle layer are wiretapped, and constraints (11) to (16) represent the conditional independence between the signals in the first layer and those in the middle layer. In particular, (16) shows that $X \rightarrow (Z_1, Z_2, Z_3) \rightarrow (S_1, \dots, S_5)$ forms a Markov chain. Note that constraints (3) to (5) and (11) to (16) implicitly allow some randomness to be injected at the corresponding nodes. We use the Xitip program [11], which relies on the framework in [13], to show that $H(X) \leq 5/3$ is implied by the set of equalities (4). Therefore, $5/3$ is an upper bound on the secrecy rate when the location of wiretapper is unknown, which is less than the secrecy rate 2 achievable when such information is known.

B. *Unequal Link Capacities (Scenario 2)*

We next show that the unachievability of the cut-set bound also holds for the secure network coding problem with unequal link capacities (Scenario 2). We convert the example of Fig. 1 by partitioning each non-middle layer link into $\frac{1}{\epsilon}$ parallel small links each of which has capacity ϵ. Any three links can be wiretapped in the transformed graph.

For the case where the location of the wiretap links is known, deleting any k' ($k' \leq 3$) non-middle layer links reduces the max flow by at most $k'\epsilon$. When $k' = 0$, the min-cut is 2. When $k' \geq 1$ or at most 2 middle layer links are wiretapped, the min-cut between the source and the sink is at least 2 after deleting these wiretapped middle layer links, and the min-cut is at least $2 - k'\epsilon \geq 2 - 3\epsilon$ after further deleting the $k' \geq 1$ non-middle layer links. Therefore, the cut-set bound is at least $2 - 3\epsilon$.

For the case where the location of the wiretap links is unknown, we prove the unachievability of the cut-set bound in the transformed network. First, consider the transformed network with the restriction that the wiretapper can only wiretap any 3 links in the middle layer. The optimal solution is exactly the same as for the original network of the previous subsection, and achieves secrecy rate at most $5/3$. Now, consider the transformed network without the restriction on wiretapping set, i.e., the wiretapper can wiretap any 3 links in the entire network. As wiretapping only the middle layer links is a subset of all possible strategies that the wiretapper can have, the secrecy rate in the transformed network is less than or equal to that in the former case, which is strictly smaller than the cut-set bound for ϵ strictly smaller than $\frac{1}{9}$. Therefore, the cut-set bound is still unachievable when the wiretap links are unrestricted in the transformed graph.



(a) Original Graph $\mathcal{H}$

(b) Transformed Graph $\mathcal{G}^{\mathcal{H}}$

Fig. 2. Example of NP-hardness proof for the case with knowledge of the wiretapping set.

IV. NP-HARDNESS

We show in the following that determining the secrecy capacity is NP-hard by reduction from the clique problem, which determines whether a graph contains a clique[1] of at least a given size r.

From Section III, finding the secrecy capacity when the location of the wiretap links is known to the communicating nodes is the same as the NP-hard network interdiction problem [14], which is to minimize the maximum flow of the network when a given number of links in the network is removed. To show that the case where the location of the wiretap links is unknown is NP-hard, we use the construction in [14] showing that for any clique problem on a given graph $\mathcal{H}$, there exists a corresponding network $\mathcal{G}^{\mathcal{H}}$ whose secrecy capacity is r when the location of the wiretap links is known if and only if $\mathcal{H}$ contains a clique of size r. We then show that for all such networks $\mathcal{G}^{\mathcal{H}}$, the secrecy capacity for the case when the location of the wiretap links is unknown is equal to that for the case when such information is known.

We briefly describe the approach in [14] in the following. Given an undirected graph $\mathcal{H} = (\mathcal{V}_h, \mathcal{E}_h)$, we will define a capacitated directed network $\hat{\mathcal{G}}^{\mathcal{H}}$ such that there exists a set of links $\hat{\mathcal{A}}'$ in $\hat{\mathcal{G}}^{\mathcal{H}}$ containing less than or equal to $|\mathcal{E}_h| - \binom{r}{2}$ links such that $\hat{\mathcal{G}}^{\mathcal{H}} - \hat{\mathcal{A}}'$ has a maximum flow of r if and only if $\mathcal{H}$ contains a clique of size r. For a given undirected graph $\mathcal{H} = (\mathcal{V}_h, \mathcal{E}_h)$ without parallel edges and self loops, we create a capacitated, directed graph $\mathcal{G}^{\mathcal{H}} = (\mathcal{N}, \mathcal{A})$ as follows: For each edge $e \in \mathcal{E}_h$ create a node i_e in a node set $\mathcal{N}_1$ and for each vertex $v \in \mathcal{V}_h$ create a node j_v in a node set $\mathcal{N}_2$. In addition, create source node s and destination node d. For each edge $e \in \mathcal{E}_h$, direct an arc in $\mathcal{G}^{\mathcal{H}}$ from s to i_e with capacity 2 and call this set of arcs $\mathcal{A}_1$. For each edge $e = (u, v) \in \mathcal{E}_h$, direct two arcs in $\mathcal{G}^{\mathcal{H}}$ from i_e to j_v and j_u with capacity 1, respectively and call this set of arcs $\mathcal{A}_2$. For each vertex $v \in \mathcal{V}_h$, direct an arc with capacity 1 from j_v to d. Let this be the set of arcs $\mathcal{A}_3$. This completes the construction of $\mathcal{G}^{\mathcal{H}} = (\mathcal{N}, \mathcal{A}) = (\{s\} \cup \{d\} \cup \mathcal{N}_1 \cup \mathcal{N}_2, \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3)$. In Fig. 2, we give an example of the graph transformation, where $\mathcal{H} = (\{1, 2, 3, 4\}, \{a, b, c, x, y\})$. We replicate [14, Lemma 2] as follows.

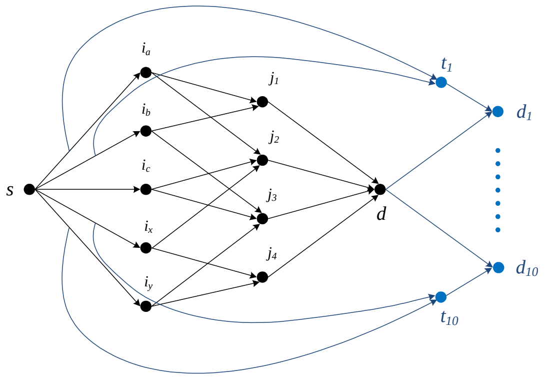


Fig. 3. Illustration of Strategy 1. In this figure, $k = 2$ and only the 5 links in the first layer can be wiretapped.

Lemma 1. *Let $\mathcal{G}^{\mathcal{H}}$ be constructed from $\mathcal{H}$ as above. Then, there exists a set of arcs $\mathcal{A}'_1 \subseteq \mathcal{A}_1$ with $|\mathcal{A}'_1| = |\mathcal{E}_h| - \binom{r}{2}$ such that the maximum flow from s to d in $\mathcal{G}^{\mathcal{H}} - \mathcal{A}'_1$ is r if and only if $\mathcal{H}$ contains a clique of size r.*

After obtaining $\mathcal{G}^{\mathcal{H}}$, we generate $\hat{\mathcal{G}}^{\mathcal{H}}$ by replacing each arc (i_e, j_v) with $|\mathcal{E}_h|$ parallel arcs each with capacity $1/|\mathcal{E}_h|$ and call this arc set $\hat{\mathcal{A}}_2$. We carry out the same procedure for arcs (j_v, d) and call this arc set $\hat{\mathcal{A}}_3$. Then $\hat{\mathcal{G}}^{\mathcal{H}} = (\mathcal{N}, \mathcal{A}) = (\{s\} \cup \{d\} \cup \mathcal{N}_1 \cup \mathcal{N}_2, \mathcal{A}_1 \cup \hat{\mathcal{A}}_2 \cup \hat{\mathcal{A}}_3)$. For the case when the location of wiretap links is known, it is shown in [14] that the worst case wiretapping set $\hat{\mathcal{A}}'$ must be a subset of $\mathcal{A}_1$. By using Lemma 1, this case is NP-hard.

Now, we consider the case where the wiretapping set is unknown, and show that the secrecy capacity of $\hat{\mathcal{G}}^{\mathcal{H}}$ when the wiretapper accesses any unknown subset of $k = |\mathcal{E}_h| - \binom{r}{2}$ links is r if and only if $\mathcal{H}$ contains a clique of size r. We use the following achievability result from [15], which uses a strategy where random keys injected by the source are either canceled at intermediate nodes or decoded by the sink:

Strategy 1 achievable rate: Connect each subset of links $\mathcal{A} \in \mathcal{W}$ in the network $\mathcal{G}$ to a virtual node $t^{\mathcal{A}}$, and connect both $t^{\mathcal{A}}$ and the actual sink to a virtual sink $d^{\mathcal{A}}$. Let $R_{s \to \mathcal{A}}$ be the minimum cut capacity between s and $t^{\mathcal{A}}$. The virtual link between $t^{\mathcal{A}}$ and $d^{\mathcal{A}}$ has capacity $R_{s \to \mathcal{A}}$, and the virtual link between the actual sink and $d^{\mathcal{A}}$ has capacity R_s. This is illustrated in Fig. 3. If the min-cut between the source and each virtual receiver $d^{\mathcal{A}}$ is at least $R_s + R_{s \to \mathcal{A}}$, the secrecy rate R_s is achievable.

From Lemma 1, the condition that $\mathcal{H}$ contains a clique of size r is equivalent to the condition that the max-flow to the sink in $\mathcal{G}^{\mathcal{H}}$ after removing any k links from $\mathcal{A}_1$ is r. We now show that the latter condition is equivalent to the condition

[1] A clique in a graph is a set of pairwise adjacent vertices, or in other words, an induced subgraph which is a complete graph.

that the secrecy capacity of $\mathcal{G}^{\mathcal{H}}$ when the wiretapper accesses any unknown subset of k links from $\mathcal{A}_1$ is r. We create a virtual sink connecting each subset of k links from $\mathcal{A}_1$ and the actual sink. As the wiretapped links are connected to the source directly, the min-cut between each virtual sink and the source is at least $2k + r$. Since r is the cut-set upper bound on the secrecy rate, by using Strategy 1 the secrecy rate r is achievable, which is equal to the secrecy rate when the location of wiretap links is known.

Finally, we show that the secrecy capacity of $\mathcal{G}^{\mathcal{H}}$ when any k links of $\mathcal{A}_1$ are wiretapped (scenario 1) is equal to the secrecy capacity of $\hat{\mathcal{G}}^{\mathcal{H}}$ when any k links are wiretapped (scenario 2). Since each second layer link has a single first layer link as its only input, wiretapping a second layer link yields no more information to the wiretapper than wiretapping a first layer link. When some links in the third layer are wiretapped, let the wiretapping set be $\hat{\mathcal{A}}' = \hat{\mathcal{A}}'_1 \cup \hat{\mathcal{A}}'_3$ where $|\hat{\mathcal{A}}'_3| \geq 1$ and $|\hat{\mathcal{A}}'_1| \leq k - 1$. Thus $\mathcal{A}_1 - \hat{\mathcal{A}}'_1$ contains at least $\binom{r}{2} + 1$ arcs. We create nodes $t^{\hat{\mathcal{A}}'}$ $d^{\hat{\mathcal{A}}'}$ with their corresponding incident links as described in Strategy 1. As removing links in $\mathcal{A}_1$ is equivalent to removing links in $\mathcal{H}$, after removing links in $\mathcal{H}$ corresponding to $\hat{\mathcal{A}}'_1$, $\mathcal{H}$ contains a subgraph $\mathcal{H}_1$ containing $\binom{r}{2}$ edges plus at least an edge $e = (u, v)$.

Case 1: $\mathcal{H}_1$ is a clique of size r. In this case, the number of vertices with degree greater than 0 in $\mathcal{H}_1 \cup e$ is $r + 2$.

Case 2: $\mathcal{H}_1$ is not a clique. $\mathcal{H}_1$ contains at least $r+1$ vertices with degree greater than 0.

According to [14, Lemma 1], the max-flow in $\mathcal{G}^{\mathcal{H}}$ is equal to the number of vertices in $\mathcal{H}$ with degree greater than 0. In both cases, the max-flow of $\mathcal{G}^{\mathcal{H}}$ after removing links in $\hat{\mathcal{A}}'_1$ is at least $r + 1$. Let $\tilde{R}_{s \to \hat{\mathcal{A}}'_3}$ be the max-flow capacity from the source to $\hat{\mathcal{A}}'_3$ in $\mathcal{G}^{\mathcal{H}} - \hat{\mathcal{A}}'_1$.

We can use a variant of the Ford-Fulkerson (augmenting paths) algorithm, e.g., [16], as follows to construct a max-flow subgraph $\mathcal{D}$ from s to $\hat{\mathcal{A}}'_3$ in $\mathcal{G}^{\mathcal{H}} - \hat{\mathcal{A}}'_1$ satisfying the property that after removing $\mathcal{D}$ from $\mathcal{G}^{\mathcal{H}} - \hat{\mathcal{A}}'_1$, the min-cut between s and d is at least

$$\begin{aligned} r + 1 - \tilde{R}_{s \to \hat{\mathcal{A}}'_3} &\geq r + 1 - |\hat{\mathcal{A}}'_3| / |\mathcal{E}_h| \\ &\geq r + 1 - (|\mathcal{E}_h| - 1) / |\mathcal{E}_h| \\ &> r, \end{aligned} \tag{5}$$

where we have used $|\hat{\mathcal{A}}'_3| \leq |\mathcal{E}_h| - 1$. Considering the network $\mathcal{G}^{\mathcal{H}} - \hat{\mathcal{A}}'_1$ with all link directions reversed, we construct augmenting paths via depth first search from d to s, starting first by constructing augmenting paths via links in $\hat{\mathcal{A}}'_3$, until we obtain a set of paths corresponding to a max flow of capacity $\tilde{R}_{s \to \hat{\mathcal{A}}'_3}$ between s and $\hat{\mathcal{A}}'_3$. We add further augmenting paths until we obtain a max flow (of capacity at least $r+1$) between s and d, which may cause some of the paths traversing links in $\hat{\mathcal{A}}'_3$ to be redefined while not changing their total capacity. The subgraph $\mathcal{D}$ consists of the final set of paths traversing links in $\hat{\mathcal{A}}'_3$. Thus, the paths remaining after removing $\mathcal{D}$ have a total capacity lower bounded by (5).

Therefore, the min-cut between the source and $d^{\hat{\mathcal{A}}'}$ in $\mathcal{G}^{\mathcal{H}} - \hat{\mathcal{A}}'_1 - \mathcal{D}$ is at least r, and the min-cut between the source and $d^{\hat{\mathcal{A}}'}$ in $\mathcal{G}^{\mathcal{H}}$ is at least $r + R_{s \to \hat{\mathcal{A}}'_1} + \tilde{R}_{s \to \hat{\mathcal{A}}'_3} = r + R_{s \to \hat{\mathcal{A}}'}$. By using Strategy 1, a secure rate of r is achievable when $\hat{\mathcal{A}}'$ is wiretapped. Thus, the secrecy rate for the case when the location of the wiretap links is unknown is equal to that for the case when such information is known with an unrestricted wiretapping set. We have thus proved the following theorem.

Theorem 2. *For a single-source single-sink network consisting of point-to-point links and an unknown wiretapping set, computing the secrecy capacity is NP-hard.*

V. Conclusion

We have considered secure network coding in the presence of a wiretapper. In contrast to previous results for networks with equal capacity links and unrestricted wiretapping sets for which the cut-set bound is identical to the secrecy capacity, we have shown for a simple example network that the secrecy capacity is strictly smaller than the cut-set bound if the network has equal capacity links and the wiretapper has only access to a restricted wiretapping set. In addition, based on this result we have also shown that the cut-set bound is not achievable in general if the wiretapping set is unrestricted but the network consists of links of unequal capacity. Finally, we have addressed the complexity of determining the secrecy capacity if the location of the wiretapper is unknown. It is shown that this problem, which is closely related to network interdiction, is NP-hard.

References

[1] A. Wyner, "The wire-tap channel," *Bell Systems Technical Journal*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[2] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst.Tech. J.*, vol. 28, pp. 656–715, 1948.

[3] N. Cai and R. Yeung, "Secure network coding," in *Proc. of IEEE ISIT*, June 2002, p. 323.

[4] J. Feldman, T. Malkin, R. Servedio, and C. Stein, "On the capacity of secure network coding," in *Proc. of Allerton Conference on Communication, Control, and Computing*, Sept. 2004.

[5] S. Y. El Rouayheb and E. Soljanin, "On wiretap networks II," in *Proc. of IEEE ISIT*, Nice, France, June 2007, pp. 551–555.

[6] N. Cai and R. W. Yeung, "A security condition for multi-source linear network coding," in *Proc. of IEEE ISIT*, Nice, France, June 2007, pp. 561–565.

[7] Z. Zhang and R. Yeung, "A general security condition for multi-source linear network coding," in *Proc. of IEEE ISIT*, Seoul, Korea, June 2009, pp. 1135–1158.

[8] A. Mills, B. Smith, T. Clancy, E. Soljanin, and S. Vishwanath, "On secure communication over wireless erasure networks," in *Proc. of IEEE ISIT*, July 2008, pp. 161–165.

[9] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Networking*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[10] T. Ho, M. Médard, J. Shi, M. Effros, and D. R. Karger, "On randomized network coding," in *Proc. of Allerton Conference on Communication, Control, and Computing*, Sept. 2003.

[11] "Xitip - information theoretic inequalities prover," http://xitip.epfl.ch/.

[12] T. Cui, "Coding for wireless broadcast and network secrecy," Ph.D. dissertation, California Institute of Technology, 2009.

[13] R. W. Yeung, "A framework for linear information inequalities," *IEEE Trans. Inform. Theory*, vol. 43, no. 6, pp. 1924–1934, Nov. 1997.

[14] R. K. Wood, "Deterministic network interdiction," *Mathematical and Computer Modeling*, vol. 17, no. 2, pp. 1–18, 1993.

[15] T. Cui, T. Ho, and J. Kliewer, "Achievable strategies for secure network coding for general networks," in *Information Theory and Applications Workshop*, 2010.

[16] B. C. Dean, M. X. Goemans, and N. Immorlica, "Finite termination of "augmenting path" algorithms in the presence of irrational problem data," *Lecture Notes in Computer Science*, pp. 268–279, 2006.