

Link-Layer Encryption Effect on Achievable Capacity in Wireless Network Coding

Claude Castellucia
INRIA
ccastel@inria.fr

Karim El Defrawy *
University of California, Irvine
keldefra@uci.edu

Gene Tsudik
University of California, Irvine
gts@ics.uci.edu

Abstract—In recent years, network coding has been enthusiastically promoted for certain wireless settings as a means of improving throughput and achieving higher capacity. Naturally, such proposals focused on communication issues, while paying less attention to security implications. This paper considers the effect of link-layer encryption (LLE) on network coding in wireless networks. LLE is a feature in current wireless networking standards (such as IEEE 802.11i) and its purpose is to mitigate traffic analysis attacks that are possible even in the presence of end-to-end encryption. We model single-source multicast using wireless network coding and show that, unfortunately, LLE significantly decreases the achievable capacity. Our goal is to understand the bounds of network coding and explore trade-offs between better security through LLE and greater capacity through network coding.

I. INTRODUCTION

In traditional networking, relaying nodes (routers) store and route individual packets. The advent of network coding dramatically changed this paradigm by requiring relaying nodes to combine data packets and transmit linear combinations thereof. As shown in [4], network coding can help achieve the maximum attainable throughput in certain networks (characterized by the max-flow min-cut theory). Such throughput is generally not achievable by routing alone (without coding). Since its introduction in 2001 [4] network coding has been successfully used in several applications, e.g., peer-to-peer content distribution systems [8], Internet tomography [7] and plain wireless networks [11] and [16]. This paper focuses on applications of network coding in wireless networks. Some recent results demonstrated how network coding can significantly improve throughput. Such efforts have mainly focused on design challenges of adopting network coding and improving performance, while neglecting security and privacy issues. In general, the benefit of network coding in wireless settings relies on the ability of nodes to overhear each others' transmissions, store and record overheard packets and use them to decode other linear combinations. One fundamental assumption is the lack of link-layer encryption (LLE) between every pair of nodes (in an ad-hoc network) or between every node and the access point (in an infrastructure-based network). While there certainly are scenarios where LLE is not used, it is not the preferred mode of operation for current wireless networking standards (e.g., IEEE 802.11i). More importantly, we believe that wireless network coding (WNC) systems

should provide security and privacy guarantees that are at least close to current wireless standards. Our argument is that, barring any security weaknesses, if a certain security feature (e.g., LLE) is in a current standards then future standards should also support it. Another reason to require LLE is to minimize traffic analysis by honest-but-curious insiders using unencrypted network-layer header fields. We note that even if end-to-end encryption is used, for example IPsec in transport or tunnel mode [3], certain fields of the IP header will still be visible (more details will be presented in Section III).

We focus on the relation between LLE and WNC, and show how the former can significantly decrease the gain afforded by the latter. We define the WNC problem with LLE requirements and show how the achievable capacity is affected by LLE in the single-source multicast setting. To that end we derive lower and upper bounds on the achievable capacity with LLE. From the insight gained from our analysis we provide directions on how to use LLE in conjunction with WNC. We then highlight open research issues and directions. We stress that the goal of this work is *not* to argue that LLE and WNC cannot coexist, but rather to understand their bounds/limits and explore trade-offs between better security and higher capacity.

The rest of the paper is organized as follows: Section II contains an overview of WNC. Next, Section III discusses security and privacy mechanisms in current wireless networking standards (i.e., IEEE802.11i). Section IV presents security challenges in wireless network coding systems, defines link-layer secure WNC and presents our model and analysis results. Section V discusses open issues and future work, and Section VI summarizes related work. Finally Section VII concludes the paper.

II. OVERVIEW OF WIRELESS NETWORK CODING

We describe the usage of WNC by presenting an example of the first practical implementation – COPE¹ [11]. COPE is designed using two main principles: 1) embracing the broadcast nature of the wireless channel, and 2) employing network coding. COPE also uses the following three techniques: 1) opportunistic listening, 2) opportunistic coding and 3) learning neighbor state. These techniques allow relaying nodes to combine similarly sized packets and broadcast their linear

¹The same security issues apply to other wireless network coding systems and are not specific to COPE. We use COPE as an example because it is one of the most prominent wireless network coding systems.

*Contact Author: <http://www.ics.uci.edu/~keldefra>

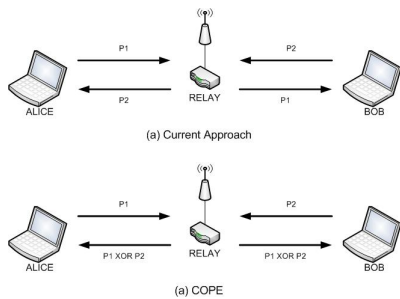


Fig. 1. A Simple Example of COPE's Operation

combinations. The set of linear combinations is constructed such that all receiving nodes can recover packets in these combinations. This, in turn, increases the overall network throughput since one transmission is used by different nodes to recover different packets. Figure 1 shows the operation of COPE where Alice and Bob send each other one packet using a wireless relay. In the plain approach Alice sends $P1$ to the relay and Bob does the same with $P2$. The relay then forwards each packet, by broadcasting it. In total four transmissions are needed to deliver both packets to their destinations. The key aspect is that the wireless medium is broadcast by nature; thus, only one transmission from the relay is needed. With WNC, the relay combines both packets (by XOR-ing them) and broadcasts the combination. Alice and Bob each know the packet they sent and can thus easily recover the other packet by simple XOR-ing. This example assumes no LLE between Alice, Bob and the relay. Whereas, if LLE is used, the network coding scenario described in Figure 1 will not work. When the relay sends a packet to Alice (or Bob) it is encrypted with Alice's (or Bob's) key and the other party cannot decrypt it. In other words, the relay can't encrypt $P1 \oplus P2$ such that both Alice and Bob can decrypt it. Possible solutions include sending $P1 \oplus P2$ unencrypted to both Alice and Bob, or to use a third key that is shared among Alice and Bob. In light of this and similar issues, we aim to understand the interplay between WNC and LLE as well as possible approaches to their coexistence.

III. SECURITY MECHANISMS IN CURRENT WIRELESS STANDARDS (IEEE802.11i)

We first briefly overview some of the security mechanisms in the IEEE802.11i standard used to secure today's wireless local-area networks. We then explain why end-to-end (E2E) encryption alone is not enough and why LLE is also required.

A. Overview of IEEE802.11i

IEEE 802.11i-2004 [2] (or 802.11i for short) is an amendment to the IEEE 802.11 standard [1] that specifies security mechanisms for wireless networks. It replaced the short authentication and privacy clause of the original standard with a detailed security clause that provides stronger alternatives for the broken wired-equivalent privacy (WEP) section. The amendment was later incorporated into the published IEEE

802.11-2007 standard. The clause describing the definition and motivation for LLE in wireless networks in 802.11i is [2]:

“ To bring the security of the wireless LAN up to the level implicit in wired LAN design, IEEE 802.11 provides the ability to protect the contents of messages. This functionality is provided by the confidentiality service.”

802.11i also introduced new key distribution techniques to overcome weaknesses in earlier methods. We highlight the following features of 802.11i:

- 1- In an independent basic service set (IBSS) network, i.e., a collection of wireless stations forming an ad-hoc network, each pair of wireless stations share a unicast key used to encrypt link-layer traffic between these two stations.
- 2- In an IBSS consisting of N nodes there are $N(N - 1)$ unicast keys to be established. Note that IBSS does not allow relaying of link-layer traffic, i.e., all nodes are one hop apart.
- 3- Each wireless station in an IBSS can have multicast and broadcast keys. Other stations in the multicast or broadcast groups of the sending stations need these keys to be able to decrypt traffic.
- 4- The setting of one access point (AP) connected to a wired network and a set of wireless stations is called basic service set (BSS). In a BSS, wireless stations can negotiate a unicast key with the AP as well as establish multicast and broadcast keys for these multicast and broadcast groups.
- 5- When a station leaves the BSS, multicast and broadcast keys need to be updated to prevent that station from decrypting subsequent traffic.

Given these features, we observe that some of them contradict the basic WNC assumption that stations can overhear each other. We note that the standard already contains mechanisms to establish unicast, multicast and broadcast keys in both infrastructure and ad-hoc modes of operation.

B. Why End-to-End Encryption is not Enough?

We recall that even if end-to-end encryption is used, for example IPsec in transport or tunnel mode [3], certain fields of the IP header will still be visible (e.g., version, header length, type of service, packet length plus authentication header size, ID, flags, fragment offset, TTL, protocol, header checksum, source and destination IP addresses). These fields can be used by honest-but-curious insider nodes to analyze the traffic patterns in a wireless network. For example source and destination IP addresses will reveal which nodes are communicating with each other and how frequent. If ToS is set, it can indicate the kind of traffic, TTL can indicate how far nodes are. The best way to prevent traffic analysis is to secure such fields using LLE. In this case an insider node not involved in the coding and/or decoding of a traffic stream will not be able to use such fields in traffic analysis. Another way to view this is that it is desired to only allow the absolute minimum required number of nodes to be able to read link-layer frames to be able to utilize WNC efficiently. We describe this in more details in the next sections.

IV. LINK-LAYER ENCRYPTION IN WIRELESS NETWORK CODING

We first present a definition of the problem of using WNC with LLE in different traffic settings (we call it link-layer secure WNC). We then model and analyze the single-source multicast case in a random wireless network and discuss our results.

A. Link-Layer Secure WNC

Problem Formulation: We assume an ad-hoc wireless network modeled as a graph $G = (V, E)$, where the set of vertices V represent the wireless stations and the set of edges E represents their connectivity. Each station $v_i \in V$ has M_i neighbors. Also, each v_i has a set of groups ($GRPS_{v_i}$). Each group ($GRP_j \in GRPS_{v_i}$) contains one or more neighbors from M_i . In other words, each group contains all nodes that know the group's encryption/decryption key. All nodes in a group ($GRP_j \in GRPS_{v_i}$) can be reached with a single transmission. Note that a neighbor can be a member of multiple groups or no groups (if it will never receive any traffic from v_i). The special case when each node has only one group containing *all* of its neighbors models the typical case considered in the WNC literature. Given a set of source and sink/terminal nodes, the link-layer secure network coding problem is to find a sequence of transmissions to deliver all traffic from all sources to all sinks (and their corresponding coding points and codes) such that receivers of every transmission are only *some (or all)* members of *only one group*. Settings for traffic between sources and sink nodes include:

- i) *single-source unicast flows*: one node sends unicast flows to several sink nodes (one sink per flow).
- ii) *multiple unicast flows*: multiple source nodes, each can send several flows but each flow has only a single sink node.
- iii) *single-source multicast flows*: a single source sends several multicast flows, where each flow has several sink nodes.
- iv) *multiple multicast flows*: multiple sources send several multicast flows, each flow has several sink nodes.

B. Network Model and Analysis

We first present our network model and then the results of our analysis.

1) **Network Model:** We use the model for *random wireless networks* proposed in [5] to study the effect of LLE on WNC. We derive bounds similar to those in [5], but when link-layer encryption is considered. In this model, two nodes u and v may not be connected even if the distance between them is less than their respective coverage ranges.

Definition 1: Quasi-Random Geometric Graph Model (G^{QRGG}): Let V be a set of n nodes selected independently, uniformly and at random from a unit square area $[0, 1]^2$. Let r and r' be two real numbers, such that $0 \leq r < r' \leq 1$. Further assume that, if two nodes are within transmission range, the probability that they are in each other's link-layer encryption groups is p_{ll} . For any two nodes u and $v \subseteq V$, we get:

- $(u, v) \in E$ if $d(u, v) \leq r$ with probability p_{ll}
 - $(u, v) \notin E$ if $d(u, v) > r'$ (outside reception range)
 - $(u, v) \in E$ with probability $p_{phy} \cdot p_{ll}$ if $r < d(u, v) \leq r'$
- p_{phy} captures the irregularity in the edges of the coverage area. The graph $G = (V, E)$ that consists of the set of vertices in V and edges in E (according to the description above) is called a **quasi-random graph** [5].

We now consider the single-source multicast setting in a random wireless network that uses WNC and LLE.

Definition 2: Connectivity Graph: Given a source s , a set of sinks T and a set of (n) relays R , we define $G_c = (V, E)$ as a graph with a set of vertices $V = s \cup R \cup T$ such that $G_c \in G^{QRGG}$. s only sends and is not connected to any nodes in T . Sink nodes only receive and are only connected to relay nodes. Relay nodes only forward traffic between source and sink nodes and can be connected either to the source, to other relay nodes, or to the sink nodes. We assume bidirectional unit capacity links (i.e., the capacity of a link $C_{ij} = C_{ji} = 1$ only if edge $(i, j) \in E$).

Definition 3: A cut in a Connectivity Graph and its Capacity: An s - t cut of size k in G_c is a partition of the n relay nodes into sets V_k and V'_k such that:

- i) $|V_k| = k$ and $|V'_k| = n - k$
- ii) $R = V_k \cup V'_k$ and $V_k \cap V'_k = \emptyset$

It follows that the total capacity of this cut is:

$$C_k = \sum_{i \in V'_k} C_{si} + \sum_{j \in V_k} \sum_{i \in V'_k} C_{ji} + \sum_{j \in V_k} C_{jt} \quad (1)$$

2) **Effect of LLE on Network Coding Capacity:** We assume that if two nodes are within each other's transmission range, they still might be unable to communicate if they are not in each other's LLE groups. To model this, we assume that the probability of two nodes being in each other's link-layer groups is uniform and constant for all nodes (p_{ll})². It follows that the probability of two nodes being connected $p_{cnct} = p_{phy} \cdot p_{ll}$, where p_{phy} is the probability of connectivity due to the locations of nodes and p_{ll} as described above is the link-layer group membership probability, p_{cnct} can thus be bounded by:

$$\frac{1}{4}(\pi r^2 + \pi(r'^2 - r^2)p_{phy}) \cdot p_{ll} \leq p_{cnct} \leq (\pi r^2 + \pi(r'^2 - r^2)p_{phy}) \cdot p_{ll} \quad (2)$$

The upper bound is a consequence of connectivity rules; it assumes that the coverage area of a node is a circle. The lower bound assumes that one of the nodes is in the corner of the unit square area (hence only a quarter of a circle). The expected value of a cut C_k can be calculated from (1) as:

$$\begin{aligned} E[C_k] &= \sum_{i \in V'_k} E[C_{si}] + \sum_{j \in V_k} \sum_{i \in V'_k} E[C_{ji}] + \sum_{j \in V_k} E[C_{jt}] \\ &= p_{cnct} \cdot (n + k(n - k)) \end{aligned} \quad (3)$$

From equation (3) we get that $\forall k$ $E[C_k] = E[C_{n-k}]$ where $0 \leq k \leq n$ and $E[C_0] = E[C_n] \leq E[C_1] = E[C_{n-1}] \leq \dots \leq E[C_{\lfloor n/2 \rfloor}]$.

²We acknowledge that this is a limiting assumption and a more representative probability distribution could be used (discussed in Section V).

Now, we need to prove that the capacity of C_k (an s - t cut) is concentrated around its expected value. Note that all edges in the connectivity graph that are incident to a common vertex are independent because the coordinates of the vertices are selected independently and uniformly at random. Consequently, random variables in C_{ij} (where i is fixed) are independent. This is an important observation that we utilize in our analysis³. Similar to [5] we derive our bounds based on the following Lemmas:

Lemma 1 (Chernoff Bound): If X_1, \dots, X_m are independent random variables such that $Pr[X_k = 1] = p$ and $Pr[X_k = 0] = 1 - p$, and $X = \sum_{k=1}^m X_k$, then for $0 < \epsilon < 1$, the following holds: $Pr[X \leq (1 - \epsilon)E[X]] \leq e^{-E[X]\epsilon^2/2}$

Proof: a proof of this well known bound is in [15].

Lemma 2: If we have an two random variables, such that $0 \leq X, Y \leq a$, then the following holds: $Pr[X + Y \leq a] \leq Pr[X \leq \frac{a}{2}] + Pr[Y \leq \frac{a}{2}]$

Proof: $Pr[X + Y \leq a] = \int_0^{a/2} \int_0^{a/2} f(x, y) dy dx + \frac{1}{2} \int_0^{a/2} \int_{a/2}^a f(x, y) dx dy + \frac{1}{2} \int_0^{a/2} \int_{a/2}^a f(x, y) dy dx$, where $f(x, y)$ is the joint pdf of X and Y . $Pr[X \leq \frac{a}{2}] = \int_0^{a/2} \int_0^{a/2} f(x, y) dy dx + \int_0^a \int_{a/2}^a f(x, y) dy dx$ and $Pr[Y \leq \frac{a}{2}] = \int_0^{a/2} \int_0^{a/2} f(x, y) dx dy + \int_{a/2}^a \int_0^{a/2} f(x, y) dx dy$. The inequality is satisfied because the right hand contains all the terms on the left hand side plus additional ones.

Theorem 1: For all cuts of size k , and all $0 < \epsilon < 1$ the following holds:

$$Pr[C_k \leq (1 - \epsilon)E[C_k]] \leq (k + 1) \cdot (e^{-\frac{\epsilon^2(n-k)p_{phy}}{2}})^{p_{ll}} \quad (4)$$

Proof: Let C_k denote the capacity of an s - t cut which contains $s \cup V_k$ on one side and $V_k' \cup t$ on the other, where s is the source node, $V_k \cup V_k'$ is a partition of the relay nodes into two disjoint sets, of cardinalities k and $n - k$ respectively. Equation (1) can be rewritten as:

$$C_k = \sum_{i \in V_k'} C_{si} + \sum_{j \in V_k} \sum_{i \in V_k' \cup t} C_{ji} \quad (5)$$

Using Lemma 2, and equation (5), if the following event happens: $C_k \leq (1 - \epsilon)E[C_k]$, then at least one of the $k + 1$ events below has to be true:

- (i) $\sum_{i \in V_k'} C_{si} \leq (1 - \epsilon)E[C_k]/(k + 1) \rightarrow$ one event
- (ii) $\sum_{i \in V_k' \cup t} C_{ji} \leq (1 - \epsilon)E[C_k]/(k + 1) \rightarrow$ k events

where $j \in V_k$. Note that values on the left side of (i) and (ii) are sums of independent random variables; thus, we can use Lemma 1 to limit the probability of these events. We thus obtain that:

$$\begin{aligned} Pr[C_k \leq (1 - \epsilon)E[C_k]] &\leq Pr[\sum_{i \in V_k'} C_{si} \leq (1 - \epsilon)E[C_k]/(k + 1)] \\ &+ \sum_{j \in V_k} Pr[\sum_{i \in V_k' \cup t} C_{ji} \leq (1 - \epsilon)E[C_k]/(k + 1)] \\ &\leq Pr[\sum_{i \in V_k'} C_{si} \leq (1 - \epsilon)(n + k(n - k))p_{cnct}/(k + 1)] \end{aligned}$$

³If nodes u and v are connected, and u is connected to yet another relay node w , there is a good chance that v is connected to w . But since the probabilities of link-layer connectivity due to group encryption keys are independent of each other, such edges will be rare and it is safe to assume that edges are independent in our connectivity graph.

$$\begin{aligned} &+ \sum_{j \in V_k} Pr[\sum_{i \in V_k' \cup t} C_{ji} \leq (1 - \epsilon)(n + k(n - k))/(k + 1)] \\ &\leq \exp(-((n + k(n - k))p_{cnct}\epsilon^2/2(k + 1))) \\ &+ k \cdot \exp(-((n + k(n - k))p_{cnct}\epsilon^2/2(k + 1))) \\ &\leq (k + 1)(\exp(-((n + k(n - k))p_{phy}\epsilon^2/2(k + 1))))^{p_{ll}} \end{aligned}$$

Note that $(n + k(n - k))/(k + 1) = (n(k + 1) - k^2)/k + 1 \approx n - k$, thus we get:

$$Pr[C_k \leq (1 - \epsilon)E[C_k]] \leq (k + 1)(\exp(-((n - k)p_{phy}\epsilon^2/2)))^{p_{ll}} \quad (6)$$

And, since $p_{ll} \leq 1$, e.g., if $p_{ll} = 1/N$, then the probability bound decreases exponentially with increase in N (note the $(k + 1)$ before the exponent).

We now show that the capacity of a minimum cut is, with high probability, concentrated around that of the source (i.e., $E[C_0]$). This indicates that the bottleneck in a random wireless network is more likely to be the connection from the source to relays. This is an important result, since it points out that there normally is enough diversity in the links between relay nodes in the middle of the network to make the bottleneck close to the source. Adding LLE between relays will not significantly influence the achievable capacity but reduces possible traffic analysis by limiting the number of honest-but-curious neighbors that could perform it. On the other hand, if the source utilizes LLE in a way that appreciably reduces the capacity of the minimum cut (the bottleneck around it), then using network coding would not be beneficial, due to insufficient link diversity.

Theorem 2: (lower and upper bounds on cut capacity)

Assume that G_c is a connectivity graph with one source node (s), a set (R) of n relay nodes and a set (T) of terminal nodes. The probability that two nodes are in each other's link-layer encryption groups is p_{ll} . The physical connectivity probability (p_{phy}) due to transmission range and locations is as in equation (2). Then, with probability $1 - O(\tau/n^2 \cdot p_{ll})$, where $\tau = |T|$, the network coding capacity $C_{s,T}$ of G_c is lower-bounded by:

$$C_{s,T} \geq (1 - \epsilon)E[C_0], \text{ where } \epsilon = \sqrt{\frac{4 \ln(n)}{p_{phy}(n-k)}}$$

and, with probability $1 - O(1/n^{4 \cdot p_{ll}/3})$, network coding capacity $C_{s,T}$ of G_c is upper-bounded by:

$$C_{s,T} \leq (1 + \epsilon)E[C_0], \text{ where } \epsilon = \sqrt{\frac{4 \ln(n^{p_{ll}})}{E[C_0]}}$$

Lower-Bound Proof: Let C_{min} be the capacity of a min s - t cut of size k (i.e., $C_{min} = C_k$). From equation (3) we see that $E[C_k] \geq E[C_0]$ for all k where $0 \leq k \leq n$. Thus, we get that: $Pr[C_{min} < (1 - \epsilon)E[C_0]] \leq Pr[C_{min} < (1 - \epsilon)E[C_k]]$

$$< 2 \exp(-(n - k)p_{phy} \cdot p_{ll}\epsilon^2/2 + \ln(k + 1))$$

By substituting ϵ with the above value, we get:

$$Pr[C_{min} < (1 - \epsilon)E[C_0]] \leq 2 \cdot \exp(-2p_{ll} \ln(n)) = O(\frac{1}{n^{2p_{ll}}})$$

Now considering all the terminal nodes in T the probability that the network coding capacity $C_{s,T}$ is below $(1 - \epsilon)E[C_0]$

can be bounded by $O(\tau/n^{2 \cdot p_{lu}})$ as follows:

$$\begin{aligned} Pr[C_{s,T} < (1 - \epsilon)E[C_0]] &\leq Pr\left[\bigcup_{t \in T} (C_{min} < (1 - \epsilon)E[C_0])\right] \\ &\leq \sum_{t \in T} Pr[(C_{min} < (1 - \epsilon)E[C_0])] = O(\tau/n^{2 \cdot p_{lu}}) \end{aligned}$$

Upper-Bound Proof: If $C_{s,T}$ exceeds $(1 + \epsilon)E[C_0]$, the capacity of any s - t cut for $t \in T$ must exceed it as well. Thus, the cut $(s; R \cup T)$ must exceed $(1 + \epsilon)E[C_0]$. Since s is not directly connected to any node in T , we obtain:

$$\begin{aligned} Pr[C_{s,T} > (1 + \epsilon)E[C_0]] &\leq Pr\left[\sum_{r \in R} C_{s,r} > (1 + \epsilon)E[C_0]\right] \\ &\leq Pr\left[\left|\sum_{r \in R} C_{s,r} - E[C_0]\right| > \epsilon E[C_0]\right] \end{aligned}$$

Note that $C_{s,r}$ are independent and identically distributed (i.i.d) random variables for all $r \in R$. We use the Chernoff bound for i.i.d random variables X_i with $Pr[X_i = 1] = p'$ [15]: $Pr\left[\left|\sum_{i=1}^n X_i - np'\right| > t\right] < 2 \cdot \exp(-t^2/3np')$ and plug in $\epsilon = \sqrt{\frac{4 \ln(n^{p_{lu}})}{E[C_0]}}$ from our hypothesis to get:

$$Pr\left[\left|\sum_{r \in R} C_{s,r} - E[C_0]\right| > \epsilon E[C_0]\right] < O(n^{-4 \cdot p_{lu}/3})$$

From the bounds we get that $3/4 \leq p_{lu} \leq 1$ for both of the bounds to be tight and the capacity of the minimum cut to be centered around that of the source with high probability.

V. OPEN ISSUES AND FUTURE WORK

We acknowledge that more analysis is needed to fully understand the interplay between LLE and WNC. To that end, several open issues remain: (1) Our analytical model could be augmented with a better abstraction of the physical channel (e.g., interference and fading effects). (2) A more detailed representation of link-layer group membership criteria is required, as we only considered group membership modeled with a fixed probability $p_{lu} \leq 1$. (3) We analyzed the single-source multicast scenario. There are other interesting cases to consider, e.g., multiple unicast flows and multiple multicast flows from many sources, as well as similar anycast scenarios.

VI. RELATED WORK

Achievable capacity and gains from network coding and code construction were studied in [4], [5], [12], [13] and [9]. Security obtainable from codes has been studied in [6], wireless systems that implement network coding were demonstrated in [11] and [16]. None of the above has looked at the interaction between LLE and WNC. Our future work will use (and build upon) some of the analysis and results of achievable capacities in wireless settings while augmenting it with our new constraints. [10] and [14] survey attacks that aim to disrupt the data delivery process in WNC systems (e.g., packet reception information mis-reporting, link-state pollution, neighbor set pollution, ACK injection or modification, packet pollution). Such attacks are out of scope of this work; but we note that some pollution attacks will be harder to execute since the adversary (even an insider) would not have access to all wireless links.

VII. CONCLUSION

In this paper, we discussed the interplay between link-layer encryption (LLE) and wireless network coding (WNC). Our main motivation is to prevent traffic analysis attacks by honest-but-curious insiders that are still possible even if end-to-end encryption is used. Additionally, current wireless standards recommend using LLE for access control and to guarantee the same level of confidentiality as in wired networks which we also expect from WNC systems. We analyzed the single-source multicast setting in a random wireless network and showed how the achievable capacity (due to WNC) drops when LLE is used and derived upper and lower bounds on such a capacity. We also showed that the bottleneck of the achievable capacity is with high probability near the source which suggests that using LLE between relays will improve security without limiting the achievable capacity from WNC.

ACKNOWLEDGMENT

The authors would like to thank Scott Jordan, Ivan Martinovic and Yanbin Lu for helpful feedback and discussions about the problem.

REFERENCES

- [1] IEEE 802.11-2007: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.
- [2] IEEE 802.11i-2004: Amendment 6: Medium Access Control (MAC) Security Enhancements.
- [3] Security Architecture for the Internet Protocol, IETF RFC 4301. <http://tools.ietf.org/html/rfc4301>.
- [4] R. Ahlswede, Ning Cai, S.-Y.R. Li, and R.W. Yeung. Network information flow. *Information Theory, IEEE Transactions on*, 46(4), Jul 2000.
- [5] S. Aly, V. Kapoor, and J. Meng. Bounds on the network coding capacity for wireless random networks. In *In Proc. 3rd Workshop on Network Coding, Theory, and Applications*, 2007.
- [6] J. Feldman, T. Malkin, C. Stein, and RA Servedio. On the capacity of secure network coding. *Proc. 42nd Annual Allerton Conference on Communication, Control, and Computing*, 2004.
- [7] C. Fragouli and A. Markopoulou. A Network Coding Approach to Network Monitoring (Invited Paper). In *43rd Annual Allerton Conference on Communication, Control, and Computing*, 2005.
- [8] C. Gkantsidis and P. R. Rodriguez. Network coding for large scale content distribution. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, volume 4, pages 2235–2245, March 2005.
- [9] Ashish Goel and Sanjeev Khanna. On the network coding advantage for wireless multicast in euclidean space. In *IPSN '08: Proceedings of the 7th international conference on Information processing in sensor networks*, pages 64–69, Washington, DC, USA, 2008. IEEE Computer Society.
- [10] J., R. Curtmola, and C. Nita-Rotaru. Secure network coding for wireless mesh networks: Threats, challenges, and directions. *Computer Communications (Elsevier)*, 32, November 2009.
- [11] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft. Xors in the air: practical wireless network coding. *SIGCOMM Comput. Commun. Rev.*, 36(4):243–254, 2006.
- [12] A. Keshavarz-Haddad and R. Riedi. Bounds on the benefit of network coding: Throughput and energy saving in wireless networks. In *INFOCOM*, 2008.
- [13] Zongpeng Li, Baochun Li, Dan Jiang, and Lap Chi Lau. On achieving optimal throughput with network coding. In *INFOCOM*, pages 2184–2194, 2005.
- [14] L. Lima, J. Vilela, P. Oliveira, and J. Barros. Network coding security: Attacks and countermeasures. *CoRR*, abs/0809.1366, 2008.
- [15] M. Mitzenmacher and E. Upfal. Probability and computing: Randomized algorithms and probabilistic analysis. Cambridge University Press, 2005.
- [16] X. Zhang and B. Li. Dice: a game theoretic framework for wireless multipath network coding. In *MobiHoc '08*.