NETWORKING WITH SECRECY CONSTRAINTS

Parvathinathan Venkitasubramaniam, Ting He and Lang Tong[†]

School of Electrical and Computer Engineering
Cornell University
Email: {pv45, th255, lt35}@cornell.edu

ABSTRACT

Wireless Networks are susceptible to anonymous monitoring of transmissions by adversaries who can infer valuable information about data flows in the network. It is therefore necessary to design network protocols that maintain secrecy of routes from eavesdroppers. In this work, we present a mathematical formulation of route secrecy when eavesdroppers observe transmission epochs of nodes. We consider networks where the nodes use receiver directed signaling schemes and each node has a strict delay constraint for relaying packets. We propose a scheduling technique to provide complete secrecy of routes, and based on that, characterize achievable rate regions for two-hop data routes under the given constraints. Furthermore, we extend the results when an additional constraint on packet loss is imposed.

Index Terms - Network Security, Traffic Mix, Scheduling, Packet Loss.

1. INTRODUCTION

Providing security is crucial to military wireless network operation. The wireless medium makes networks vulnerable to a wide range of attacks by adversaries. Active attacks such as jamming or node replication are countered by using sophisticated intrusion detection mechanisms [1]. Passive attacks such as traffic analysis or flow correlation attacks, wherein eavesdroppers monitor transmissions from nodes, are, however, not detectable. Passive traffic monitoring can provide adversaries with mission critical information including source-destination pairs and routes used in the network. It is therefore necessary to design secure network protocols such that the routes of information flow in the network are undetectable to eavesdroppers monitoring the node transmissions.

The strategies adopted to prevent traffic analysis attacks are dependent on the types of information available to the eavesdropper. By using encryption and packet padding, it is possible to prevent analysis based on the contents and lengths of packets (ex: Chaum's Mix-net [2]). However, by merely correlating the time points of transmissions from multiple nodes, an adversary can infer the routes of information flow, especially in low latency networks. Therefore, a key aspect in providing route secrecy is the design of transmission schedules so that correlation of transmission epochs reveals minimal information about data flows.

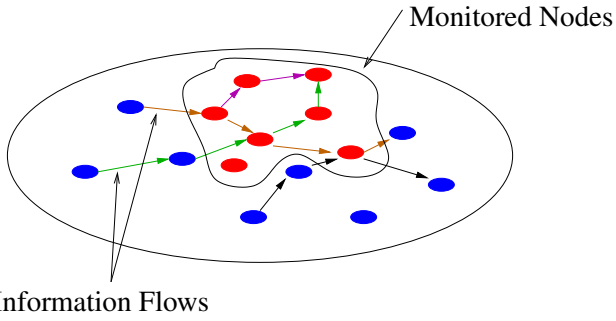

Fig. 1: Wireless Network with Eavesdroppers

As an example, if nodes always transmit packets at fixed epochs irrespective of the routes of data flow, then it is impossible to detect traffic flows by correlating packet departure times. Maintaining a fixed transmission schedule, however, would increase end-to-end delay and require transmission of dummy packets thereby reducing the network efficiency [3]. Hence, it is necessary to maximize the achievable network performance when providing route secrecy.

In this work, we consider the problem of hiding information flows when the eavesdropper has access to transmission epochs and is aware of the transmission strategy. We propose a mathematical formulation of route secrecy with respect to transmission epoch monitoring. We propose a solution to obtain complete secrecy and characterize achievable data rates for a multiplex relay (see Fig. 2) with receiver directed signaling when there is a fixed delay constraint on relayed packets. We also extend the results when an additional constraint on packet loss is imposed.

1.1. Related Work

Designing countermeasures to traffic analysis attacks is a classical problem. Many solutions [4] have been derived

[†]This work is supported in part by TRUST (The Team for Research in Ubiquitous Secure Technology), which receives support from the National Science Foundation (NSF award number CCF-0424422) and the following organizations: Cisco, ESCHER, HP, IBM, Intel, Microsoft, ORNL, Qualcomm, Pirelli, Sun and Symantec, and the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

from Chaum's Mix-net concept [2]. A Mix relays data for multiple flows and by reordering and re-encrypting the data, flow correlation of incoming and outgoing data is prevented. The idea has been used effectively in providing anonymous communication for internet applications [5–8].

For low latency networks, it has been shown in [9] that simple mixing techniques are not effective to prevent correlation of transmission epochs. They propose the use of dummy packets to make departure epochs identical irrespective of the flows. The idea of having fixed transmission schedules independent of routes has also been considered in [3, 9], where the authors give bounds on the performance loss incurred due to the secrecy constraints. In [10], the authors use randomized transmissions to prevent flow correlation in networks without latency constraints. The use of randomized routes as a countermeasure to traffic analysis attacks has also been considered in [11, 12].

The theoretical framework for secrecy in this work is motivated by the notion of equivocation developed by Shannon in [13]. The secrecy constraint we consider is a special case of Shannon's equivocation, known as maximum secrecy [14], wherein the observations provide zero information about the source.

2. PROBLEM SETUP

2.1. Definitions

Let the network be represented by a directed graph $G = (V, E)$, where V is the set of nodes and E is the set of links between pairs of nodes. A link (A, B) belonging to E denotes that node B can listen to the transmissions from A. Let $\mathcal{Y}_A = \{Y_A(1), Y_A(2), \cdots\}$ denote the time instants (known as *departure epochs*) at which A transmits packets. The *transmission rate* T_A of a node A is defined as the average number of packets per unit time transmitted by A. In other words,

$$T_A = \lim_{n\to\infty} \frac{n}{Y_A(n)}.$$

In this work, we propose techniques to hide the presence of a one-hop relay from an eavesdropper. In general, the tasks carried out by a relay can be multivarious; it can choose to decode and re-encode blocks of packets, it can relay unaltered packets after a random delay or it can reorder the packets before transmission. Re-encryption and packet padding occur at every node/relay to prevent any content based correlation. We are concerned with the kind of traffic, wherein each packet needs to be relayed within a fixed delay constraint Δ. We restrict the tasks of a relay to packet-reordering and timing perturbation. Depending on its transmission schedule, a relay picks departure epochs for the arriving packets such that the delay constraint is satisfied. Any packet that is not relayed within Δ time units after arrival is dropped. A formal definition of the relay function is given as follows.

Let $\mathcal{Y}_A = \{Y_A(1), Y_A(2), \cdots, Y_A(n)\}$ represent the departure epochs of packets from node A and let $\mathcal{Y}_B = \{Y_B(1), Y_B(2), \cdots, Y_B(n)\}$ represent the departure epochs of packets from node B. A 1×1 *relay map* is an algorithm that picks a subsequence $\mathcal{Y}_A^s$ of $\mathcal{Y}_A$ and an equal length subsequence $\mathcal{Y}_B^s$ of $\mathcal{Y}_B$ such that $\forall i, 0 \le Y_B^s(i) - Y_A^s(i) \le \Delta$.

If $|\mathcal{Y}_A| = n$ and $|\mathcal{Y}_A^s| = k(n)$, then the *relay rate* $\lambda(\mathcal{M})$ of the 1×1 relay map $\mathcal{M}$ is given by

$$\lambda = \lim_{n\to\infty} \frac{k(n)}{Y_A^s(k(n))}.$$

The rate of a relay map is dependent on the transmission rates of the nodes.

The map for a node relaying multiple flows can be defined analogously. An $m \times 1$ relay map is an algorithm that picks subsequences $\mathcal{Y}_{A_1}^s, \mathcal{Y}_{A_2}^s, \cdots, \mathcal{Y}_{A_m}^s$ from departure epochs of m nodes $A_1, \cdots, A_m$ and a subsequence Y_B^s from the departure epoch of the relay node B such that

1. $|\mathcal{Y}_B^s| = \sum_{i=1}^m |\mathcal{Y}_{A_i}^s|$.
2. Let $\mathcal{Y}^s$ be the sequence formed by the concatenating $\mathcal{Y}_{A_1}^s, \cdots, \mathcal{Y}_{A_m}^s$ and ordering the epochs in ascending order. Then,

$$\forall i \le |\mathcal{Y}^s|, 0 \le Y_B^s(i) - Y^s(i) \le \Delta.$$

An $m \times 1$ relay map is associated with a relay rate vector $\lambda(\mathcal{M}) = (\lambda_1, \cdots, \lambda_m)$ which is given by

$$\lambda_i = \lim_{n\to\infty} \frac{k_i(n)}{Y_{A_i}^s(k_i(n))},$$

where $k_i(n) = |\mathcal{Y}_{A_i}^s|$.

2.2. Medium Access Constraints

Nodes in a wireless network share a common channel and transmissions are susceptible to fading and interference. Depending on the PHY model, the rates of transmission are subjected to some medium access constraints specified by a region of rate vectors $\mathcal{C}$. Packets can be received successfully at the destination nodes only if the transmission rates belong to the set $\mathcal{C}$. To this extent, we consider receiver directed spread spectrum signaling to characterize medium access conditions.

Receiver Directed Signaling: The nodes transmitting to a common node/relay use the same spreading sequence.

The sum-rate of nodes transmitting to a single receiver is therefore bounded by a maximum value. If a single relay B serves nodes $A_1, \cdots, A_n$, the region $\mathcal{C}$ is given by

$$\mathcal{C} = \{(T_{A_1}, \cdots, T_{A_n} : \sum T_{A_i} \leq C_B\}. \quad (1)$$

When considering traffic flows across multiple hops, the PHY layer constraints coupled with stability give rise to bounds on the actual rate of data flow. For example, the maximum packet rate for a simple two-hop system with a 1×1 relay is bounded by $\min C_B, C_D$, where C_B, C_D are bounds on the transmission rates to the relay and the destination respectively.

2.3. Secrecy

When designing secure transmission schedules, it is necessary to analytically model security or secrecy provided. Our definition of secrecy is motivated by Shannon's notion of equivocation [13], which was utilized in defining secrecy in wiretapped [14] and broadcast channels [15].

Let $\mathcal{A} = \{A_1, A_2, \cdots, A_k\} \subset V$ be a subset of nodes. We define node A to be *connected* to node B (or $A \rightarrow B$), if there exists a path from node A to node B. In other words, $A \rightarrow B$ iff there exists nodes $A_1, A_2, \cdots, A_m$ such that $(A, A_1), (A_1, A_2), \cdots, (A_n, B) \in E$. Let $\mathcal{F}$ denote the set of all connected node pairs in $\mathcal{A}$,

$$\mathcal{F}_{\mathcal{A}} = \{(A, B) : A, B \in \mathcal{A}, A \rightarrow B\}.$$

It is necessary to consider all possible node-pairs (not necessarily connected by an edge), since by correlating epochs of physically distant nodes, it may be possible to gain information about the end-to-end flow.

During a given session, the set of node-pairs in $\mathcal{F}_{\mathcal{A}}$ that require non-zero relay rate is denoted by the flow vector $F \subset \mathcal{F}$. We define $\mathcal{A}$ to have *complete relay secrecy* if the flow vector F and the transmission epochs of the nodes in $\mathcal{A}$ are independent. In other words, the conditional distribution

$$p(\mathcal{Y}_{A_1}, \mathcal{Y}_{A_2}, \cdots, \mathcal{Y}_{A_k}|F) = p(\mathcal{Y}_{A_1}, \mathcal{Y}_{A_2}, \cdots, \mathcal{Y}_{A_k}) \forall F. \quad (2)$$

During any session, the eavesdropper observes the same joint distribution of transmission epochs, hence it is impossible to infer the flow by correlating time points. Statistical independence of transmission schedules and underlying flows corresponds to the notion of maximum secrecy [14, 15].

2.4. Achievable Rates

A rate vector $\mathbf{R} = (R_1, \cdots, R_m)$ for a set of node-pairs with common relay $\{(A_1, B), (A_2, B), \cdots, (A_m, B)\}$ is an achievable rate vector, if there exists a conditional distribution $p(\mathcal{Y}_{A_1}, \mathcal{Y}_{A_2}, \cdots, \mathcal{Y}_{A_m}|F)$ and an $m \times 1$ relay map such that following conditions are satisfied

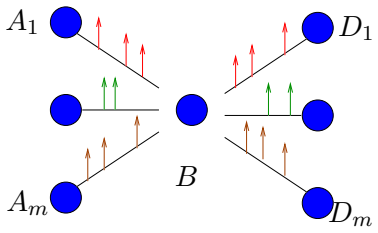

Fig. 2: Two Hop Relay

1. The transmission rate $\{T_{A_1}, T_{A_2}, \cdots, T_{A_m}, T_B\}$ satisfy the medium access constraints (1).
2. For every realization $(\mathcal{Y}_{A_1}, \cdots, \mathcal{Y}_{A_m})$,
$$\lambda_i(\mathcal{M}) \geq R_i, i = 1, \cdots, m.$$
3. $\{A_1, \cdots, A_m, B\}$ have complete relay secrecy.

In the following sections, we present achievable rate regions for the special case of providing relay secrecy for an $m \times 1$ multiplex relay (Fig. 2), where a single node relays packets from m nodes. The results are presented for the PHY model discussed in Section 2.2.

3. ACHIEVABLE RATES

In the absence of eavesdroppers, the flow-rates achievable in a network can be obtained purely from medium access and stability restrictions. In the presence of eavesdropper, however, the secrecy condition imposes additional constraints when designing transmission schedules.

The secrecy condition in (2) indicates that the distribution of transmission epochs are independent of the flows. A special case of this condition is when the transmission schedule of each node is drawn from an independent distribution and the distribution is not dependent on the flows. This notion has been considered in literature [3, 9], wherein the transmission schedules were deterministic and independent of the flows. Statistical independence of departure epochs is a sufficient condition to ensure relay secrecy. In general, it may be possible to design schedules such that the transmission epochs are not independent and yet guarantee relay secrecy.

We assume that the sources generate packets at Poisson time points which determine the schedules of the source nodes. In order to satisfy the secrecy condition, the relay nodes generate departure epochs from independent Poisson

processes. To an eavesdropper monitoring the nodes, it is impossible to decipher the actual flows by observing time points, since at all times, the transmission epochs are statistically independent. However, due to the delay constraint, the secrecy condition leads to a reduced rate region, which is characterized in the following sections.

3.1. Rate Region

Since the spreading sequences are receiver directed, the constraints on transmission rates are independent for different receiving nodes. When characterizing the achievable rates for an $m \times 1$ relay, we assume that the final destination nodes are different. Therefore, the constraint on the rates of the relay node are independent for each flow.

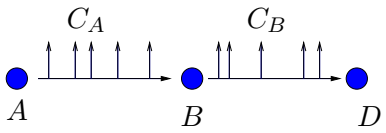

Fig. 3: 1×1 Relay

To characterize the achievable rates for a 1×1 relay map, we use the BOUNDED-GREEDY-MATCH (BGM) algorithm proposed in [16] that optimally maps Point processes with the least packet drops. Since epochs are generated according to independent Poisson processes, the delay constraint makes it impossible to relay all transmitted packets. Hence, the relay rate is strictly less than the transmission rates of the nodes.

Let node A be the transmitting node and B the relay. The algorithm is as follows; When a packet arrives at B, if there exists a departure epoch within Δ of the arrival instant and has not been matched to any previous arrival, it is assigned to the arrived packet. Otherwise, the packet is dropped. The transmission schedule of A is obtained from the generation times of packets while node B generates an independent Poisson process of a fixed rate and uses the algorithm to map arrival epochs to the generated schedule.

Theorem 1 *If the maximum transmission rates allowed to nodes B and D are C_B and $C_D$respectively, the maximum achievable relay rate R between (A, D), when $\mathcal{Y}_A, \mathcal{Y}_B$ are independent Poisson processes is obtained when $T_A = C_B, T_B = C_D$ and is given by*

$$R = \begin{cases} C_B \frac{C_D\left(e^{-\Delta(C_B - C_D)} - 1\right)}{C_D e^{-\Delta(C_B - C_D)} - C_B} & C_B \neq C_D \\ \frac{C_B^2 \Delta}{1 + C_B \Delta} & C_B = C_D \end{cases}. \quad (3)$$

Proof: Refer to Appendix.

As is evident from the expression in Theorem 1, as $\Delta \rightarrow \infty$, the relay rate approaches the best possible rate, $\min\{C_D, C_B\}$. Similarly as $C_B \rightarrow \infty$, the maximum rate is C_D for any finite Δ and vice-versa. A special case of this result, when nodes have equal transmission rates was obtained in [17] under a different context. Clearly, when Δ is finite, the transmission rates T_A, T_B of the nodes are strictly greater than the achievable information relay rate, thereby resulting in packet drops. Packet losses can, however, be countered if the source employs forward error correcting (FEC) schemes.

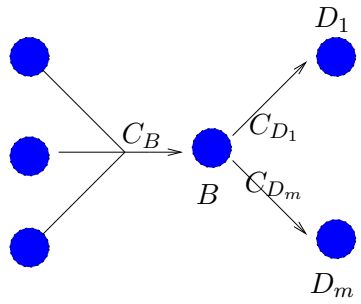

Fig. 4: Two Hop Relay

For an $m \times 1$ relay, we first consider the case when each traffic flow has a distinct destination node. Since the PHY layer is a receiver directed signaling scheme and each traffic has a unique destination node, the outgoing streams from the relay are independent point processes. Each point process is drawn from a Poisson distribution depending on the transmission rate allowed to that destination.

The relay node can decode the packet header and distinguish packets arriving from multiple sources. Although all incoming streams use the same spreading sequence, the relay can observe distinct processes from each source. The $m \times 1$ relay function therefore decouples into m 1×1 relay maps. The relay node uses the BGM algorithm to map the packets in each individual arrival process to the corresponding destination stream. Since the BGM algorithm has been proven to minimize the packet loss [17], this strategy provides the best achievable rates, when the transmission schedules are drawn from independent Poisson processes.

Theorem 2 *Let $A_i, \cdots, A_m$ be the transmitting nodes, B the relay and $C_1, \cdots, C_m$ the final destination nodes. The achievable rate region for the m flows $\mathcal{R}$ is given as follows. $(R_1, \cdots, R_m) \in \mathcal{R}$ iff $\exists\, T_{A_1}, \cdots, T_{A_m}$ such that*

$$\begin{aligned} R_i &\leq T_{A_i} \frac{C_{D_i}\left(e^{-\Delta(T_{A_i} - C_{D_i})} - 1\right)}{C_{D_i} e^{-\Delta(T_{A_i} - C_{D_i})} - T_{A_i}}, \\ \sum_i T_{A_i} &\leq C_B. \end{aligned}$$

Proof: Follows from Theorem 1.

As is evident from the theorem, if the delay constraint were infinite, the achievable region would correspond to the region without any secrecy constraints,

$$\sum_i R_i \leq C_B, \;\; R_i \leq C_{D_i}, \forall i.$$

An plot for a 2×1 relay example is shown in Figure 5.

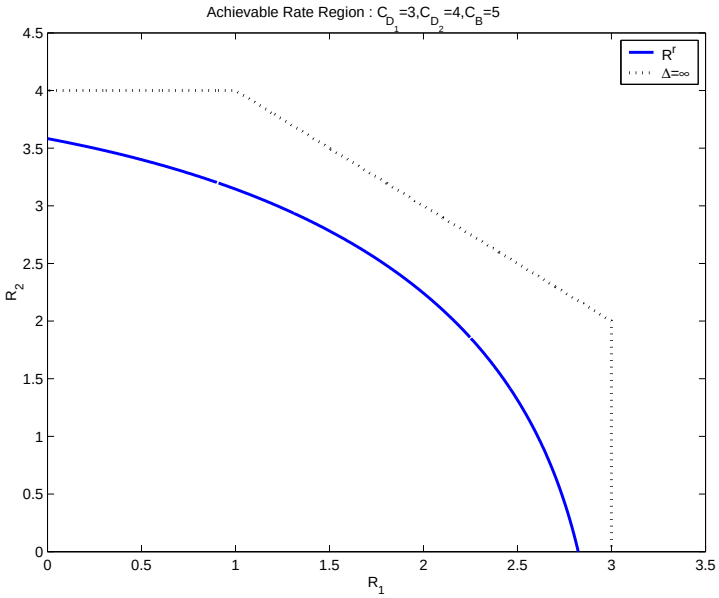

Fig. 5: Achievable Regions for 2×1 multiplex relay : $\Delta = 1$

When the destination nodes for the different traffic sources are distinct, the relay map, as mentioned earlier, decouples to individual relay maps for each traffic flow. If nodes share destinations, however, this is not possible as the outgoing streams have a sum-rate constraint for the destination.

Unlike the relay with distinct destinations, the individual input streams are mapped to the same departure process (Fig. 6). Here again, we use the optimality of the BGM algorithm in obtaining the best set of achievable rates. Although the relay can distinguish the individual input processes, the best achievable rates are obtained by using the BGM algorithm on the joint incoming process and the departure process. The rates are characterized by the following theorem.

Theorem 3 *If C_B and C_D are the maximum allowed transmission rates to the nodes B and D, then the set of achievable rates for a system shown in Figure 6 is given by*

$$\sum_i R_i \leq \begin{cases} C_B \frac{C_D\left(e^{-\Delta(C_B - C_D)} - 1\right)}{C_D e^{-\Delta(C_B - C_D)} - C_B} & C_B \neq C_D \\ \frac{C_B^2 \Delta}{1 + C_B \Delta} & C_B = C_D \end{cases}.$$

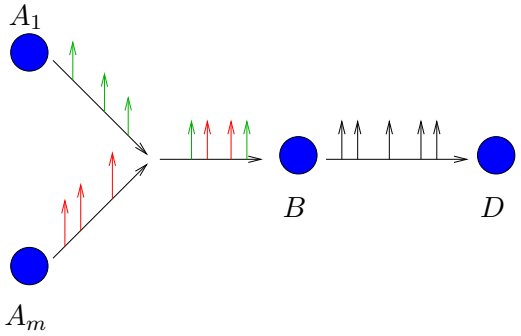

Fig. 6: Relay with shared destination

Proof: Follows from Theorem 1.

Since the sum-rates of the input and output processes are constant and the relay does not distinguish input sources, it is easy to see that the rate region is linear. By combining the techniques proposed in Theorems 2 and 3, it is possible to obtain the set of achievable rates for any arbitrary pairing of source destination pairs through a common relay.

4. PACKET LOSS CONSTRAINT

As mentioned in Section 3.1, the finite delay constraint imposed on the transmission schedule results in packet loss. Hence, it is necessary for the source to use a forward error correction scheme to ensure reliable recovery of packets at the destination. Coding for packet recovery has been addressed in literature [18, 19]. In particular, in [18], the authors propose coding schemes to recover packets when transmissions result in packet erasures.

Since packets can be appended with a sequence number, the erasure positions are known to the receiver. For a fixed block length, the information packet rate reliably delivered would be strictly less than the capacity of the erasure channel. However, as the block length of packets considered increases, it is possible to design codes with rates arbitrarily close to capacity. It can be shown that, for the relay schemes considered, the packet drop model is equivalent to a channel with stationary and ergodic erasures. Hence, as the block length increases, it is possible to obtain an end-to-end information packet rate of $1 - \epsilon$ [20], where ϵ is the fraction of packets dropped.

In practice, it may be necessary to design strategies for a fixed packet drop fraction ϵ depending on the end-to-end delay allowed and availability of good codes. The following theorem characterizes an achievable rate region for the $m \times 1$ relay (with distinct destinations), such that the packet drop fraction is always less than a fixed ϵ.

Theorem 4 *If the achievable rate region without packet loss*

constraint is given by $\mathcal{R}_r$, then the achievable relay rate region $\mathcal{R}_\epsilon$ for the 2×1 relay with packet loss constraint ϵ is given by

$$\mathcal{R}_\epsilon = \mathcal{R}_r \cap \mathcal{S}_\epsilon,$$

where

$$\mathcal{S}_\epsilon = \{(R_1, \cdots, R_m) : R_i \leq x_i(1-\epsilon),\ i = 1, \cdots, m\},$$

and x_i is the solution of

$$\epsilon = \frac{C_{D_i} - x_i}{C_{D_i} \exp(-\Delta(x_i - C_{D_i})) - x_i}. \tag{4}$$

Proof: Refer to Appendix

The rate region in Theorem 4 is obtained by using the relay map scheme described in Section3.1 coupled with the constraint on transmission rate due to the packet loss fraction ϵ.

5. CONCLUSIONS

In this work, we formally defined the problem of hiding data flows from eavesdroppers observing transmission epochs. We proposed a possible solution for providing complete secrecy and characterized achievable rates for a multiplex relay in Poisson traffic. Allowing relays to perform re-encoding is a worthwhile extension to pursue. Although we have considered only a single relay system, the basic ideas are extendable to longer routes also.

6. REFERENCES

[1] S. Axelsson, "Intrusion detection systems: A taxonomy and survey," tech. rep., Chalmers University of Technology, Sweden, March 2000.

[2] D. Chaum, "Untraceable electronic mail, return addresses and digital pseudonyms," *Communications of the ACM*, vol. 24, pp. 84–88, February 1981.

[3] B.Radosavljevic and B. Hajek, "Hiding traffic flow in communication networks," in *Military Communications Conference*, 1992.

[4] A. Serjantov and G. Danezis, "Towards an information theoretic metric for anonymity," in *Proceedings of Privacy Enhancing Technologies Workshop (PET 2002)* (R. Dingledine and P. Syverson, eds.), Springer-Verlag, LNCS 2482, April 2002.

[5] Q. Sun, D. R. Simon, Y. Wang, W. Russell, V. N. Padmanabhan, and L. Qiu, "Statistical identification of encrypted web browsing traffic," in *Proceedings of the 2002 IEEE Symposium on Security and Privacy*, (Berkeley, California), p. 19, May 2002.

[6] C. Gulcu and G. Tsudik, "Mixing e-mail with babel," in *Proceedings of the Symposium on Network and Distributed System Security*, pp. 2–19, February 1996.

[7] G. Danezis, R. Dingledine, and N. Mathewson, "Mixminion: design of a type iii anonymous remailer protocol," in *Proceedings of 2003 Symposium on Security and Privacy*, pp. 2–15, May 2003.

[8] M. K. Reiter and A. D. Rubin, "Crowds: anonymity for Web transactions," *ACM Transactions on Information and System Security*, vol. 1, no. 1, pp. 66–92, 1998.

[9] Y. Zhu, X. Fu, B. Graham, R.Bettati, and W. Zhao, "On flow correlation attacks and countermeasures in mix networks," in *Proceedings of Privacy Enhancing Technologies workshop*, May 26-28 2004.

[10] X. Hong, P. Wang, J. Kong, Q. Zheng, and J. Liu, "Effective Probabilistic Approach Protecting Sensor Traffic," in *Military Communications Conference, 2005*, (Atlantic City, NJ), pp. 1–7, Oct. 2005.

[11] S. Jiang, N. H. Vaidya, and W. Zhao, "A mix route algorithm for mix-net in wireless mobile ad hoc networks," in *Proceedings of IEEE Mobile Sensor and Ad-hoc and Sensor Systems*, pp. 406–415, October 2004.

[12] J. Kong and X. Hong, "Anodr: Anonymous on demand routing with untraceable routes for mobile ad-hoc networks," in *ACM lntemational Symposium on Mobile Ad Hoc Nefworking and Computing*, (Annapolis, MD), June 2003.

[13] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, 1949.

[14] A. Wyner, "The wiretap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–1387, 1975.

[15] I. Csiszár and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. on Information Theory*, vol. 24, pp. 339–348, May 1978.

[16] A. Blum, D. Song, and S. Venkataraman, "Detection of Interactive Stepping Stones: Algorithms and Confidence Bounds," in *Conference of Recent Advance in Intrusion Detection (RAID)*, (Sophia Antipolis, French Riviera, France), September 2004.

[17] T. He and L. Tong, "Detecting Stepping-Stone Traffic in Chaff: Fundamental Limits and Robust Algorithms," Tech. Rep. ACSP-TR-06-06-01, Cornell University, June 2006. `http://acsp.ece.cornell.edu/pubR.html`.

[18] N. Shacham and P. McKenney, "Pakcet Recovery in High-Speed Networks using Coding and Buffer Management," in *Proc. IEEE INFOCOM*, pp. 124–131, 1990.

[19] L. Rizzo, "Effective Erasure Codes for Reliable Computer Communication Protocols," in *Proc. ACM SIGCOMM Computer Communication Review*, vol. 27, pp. 24–36, 1997.

[20] S. Boucheron and M. R. Salamatian, "About priority encoding transmission," *IEEE Trans. Inform. Theory*, vol. 46, March 2000.

[21] D. Cox and H. Miller, *The Theory of Stochastic Processes*. New York: John Wiley & Sons Inc., 1965.

Appendix

Proof of Theorem 1

To prove the theorem, we adopt the technique used in [17]. Consider the two point processes $\mathcal{Y}_A, \mathcal{Y}_B$. If a packet in $\mathcal{Y}_A$, say at time t is designated as dummy packet by the BGM algorithm, we insert a virtual packet at the $t + \Delta$ in $\mathcal{Y}_B$. Similarly, if a packet at time t in $\mathcal{Y}_B$ is designated as dummy packet, we insert a virtual packet at time t in $\mathcal{Y}_A$. Now we consider the difference process $\mathcal{Z} = \{Y_B(i) - Y_A(i)\}$ between the two processes. At every occurrence of a dummy packet, the difference process hits a reflecting barrier, either at 0 or at Δ. The net probability of chaff is, therefore, the probability of hitting either barrier.

If the transmission rates of node A and B are T_A and T_B respectively, from the analysis in [21], we know that the probability of hitting Δ is given by

$$\Pr\{Z(i) = \Delta\} = \frac{1 - \frac{T_A}{T_B}}{\frac{T_B}{T_A} e^{-\Delta(T_A - T_B)} - \frac{T_A}{T_B}}.$$

It is easy to see that the fraction of chaff in $\mathcal{Y}_A$ is

$$\epsilon_A = \frac{T_B \Pr\{Z(i) = \Delta\}}{T_A(1 - \Pr\{Z(i) = \Delta\})} = \frac{T_B - T_A}{T_B e^{-\Delta(T_A - T_B)} - T_A}.$$

Since the rate of relayed packets increases with the transmission rates of either nodes, the achievability of the theorem is proved. In [16], the authors have shown that the BGM algorithm inserts the least chaff fraction for any pair of point processes. Hence, for any (T_A, T_B), it is impossible to obtain a higher information relay rate than (3).

□

Proof of Theorem 4

Fro the proof of Theorem 1, we know that the packet loss epsilon for each input-output pair of rates T_{A_i}, C_{D_i} respectively can be written as

$$\epsilon = \frac{C_{D_i} - T_{A_i}}{C_{D_i} \exp(-\Delta(T_{A_i} - C_{D_i})) - T_{A_i}},$$

when the relay transmits at the highest rate.

It is easily shown that ϵ is an increasing function of T_{A_i}. Hence, an upper bound on ϵ corresponds to an upper bound on the transmission rate $T_{A_i} + T_{A_2}$. Therefore, for any rates pair that satisfy $T_{A_i} \leq x_i$ where x_i is given by 4, the relay map guarantees that relay rates satisfy the packet loss constraint.

□

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government.