Anonymous Networking with Localized Eavesdroppers: A Game-Theoretic Approach

Parv Venkitasubramaniam and Lang Tong
School of Electrical and Computer Engineering
Cornell University, Ithaca,
NY, 14850
Email: {pv45,lt35}@cornell.edu

***Abstract*—The problem of anonymous wireless networking is considered when an adversary monitors the packet transmission timing of an unknown fraction of the network nodes. For a given level of network performance, as measured by network throughput, the problem of maximizing anonymity is studied from a game-theoretic perspective. Using conditional entropy of routes as a measure of anonymity, this problem is posed as a two player zero-sum game between the network designer and the adversary; the task of the adversary is to choose a subset of nodes to monitor so that anonymity of routes is minimum and the task of the network designer is to choose a subset of nodes (referred to as *covert relays* to generate independent transmission schedules and evade flow detection so that anonymity is maximized. It is shown that a Nash equilibrium exists for a general category of finite networks. The theory is applied to the numerical example of a switching network to study the relationship between anonymity, fraction of monitored relays and the fraction of covert relays.**

Keywords– traffic analysis, anonymity, equivocation, Nash equilibrium.

I. Introduction

A. Motivation

The packet transmission times of nodes in a network can reveal significant information about the source-destination pairs and routes of traffic flow in the network [1]. Equipped with such information, a malicious adversary can launch more powerful attacks such as jamming or denial of service. The typical design of anonymous networking protocols models adversaries as omniscient and capable of monitoring every single transmission in the network perfectly. From a practical standpoint, this is far too conservative, and such global information would be available only to the network owner or a centralized controller. In this work, our goal is to study the problem of anonymity in networks under a more general adversary model, where an *unknown* subset of the nodes are monitored by the adversary. From a network design perspective, the goal is to design transmission and relaying strategies such that the desired level of network performance is guaranteed with maximum *anonymity of network routes*. Providing anonymity to the routes of data flow in a network requires modification of packet transmission schedules, which reduces the achievable network performance. Therefore, depending on the level of network performance desired, it is necessary to pick the optimal set of nodes to modify transmission schedules so that the quality of service criterion is met while providing maximum anonymity. If the network designer were aware of which parts of the network were bring monitored by the adversary, this set of nodes can be chosen such that minimum information is available through the monitored nodes. However, if the adversary is aware of the set of nodes chosen to modify schedules, then he can choose to monitor only those nodes that provide him maximum information about the network routes. This "interplay" between the network designer and the adversary is the main subject of this work, and it is studied using a game-theoretic approach.

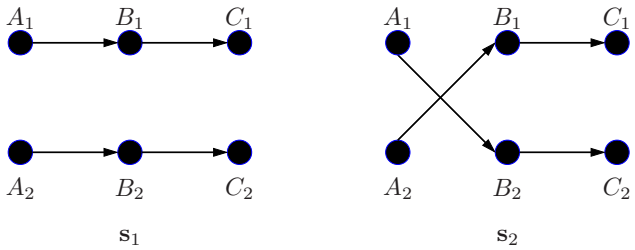

Fig. 1. $2-$relay parallel network: Two possible sessions $\mathbf{s}_1$ and $\mathbf{s}_2$.

To understand the game-theoretic perspective, consider the example of a $2-$relay parallel network as shown in Figure 1. During any period of observation of the adversary, we assume that the network operates in one of two configurations $\mathbf{s}_1$ or $\mathbf{s}_2$ (see Figure 1). The adversary's goal is to identify which of these configurations (henceforth referred to as *network session*) is currently active in the network by monitoring the time-points of transmission of all the nodes. Consider a transmitter directed signaling model, where each node transmits on a unique orthogonal channel such that transmissions of multiple nodes are non interfering. Under this signaling scheme, an adversary would have to detect correlations across transmission schedules of a source and a relay to identify the flow of traffic. Therefore, if a relay uses a transmission schedule that is statistically independent of the arrival process from the source, then the adversary would not be able to detect the flow of traffic through the relay.

Consider a scenario when the throughput requirement mandates that at most one relay can generate independent schedules (using dummy transmissions), and consider an adversary who monitors the packet transmission times of at most two

978-1-4244-2734-5/09/$25.00 ©2009 IEEE

nodes in the network. If only relay B_1 generates a transmission schedule that is statistically independent of that of A_1 and A_2, then the optimal strategy for the adversary would be to monitor (A_2, B_2) or (A_1, B_2), which would help him perfectly determine the session. However, given the knowledge that the adversary would monitor (A_1, B_2) or (A_2, B_2), the optimal strategy of the network designer would be to maximize anonymity by making the schedule of B_2 always independent. We are interested in determining if there exists a pair of strategies for the network designer and the adversary that neither has any incentive to modify. In other words, if we formulate this as a two-player zero-sum game between the adversary and the network designer with anonymity as the payoff, does a Nash equilibrium exist? As will be shown in the subsequent sections, a Nash equilibrium exists in the class of randomized strategies. By definition, at the Nash equilibrium, neither the network designer nor the adversary have any incentive to modify their strategy (See Theorem 3).

B. Main Contributions

In this work, we consider a game-theoretic formulation of anonymous networking in a general class of finite wireless networks when the number of nodes monitored by an adversary model is bounded by a known constant. We pose the design problem as a two player zero sum game with equivocation (conditional entropy) of network routes as the payoff; the adversary's strategy is to pick a random subset of nodes to monitor, and the network designer's strategy is to pick a random subset of nodes to use covert transmission schedules. For the class of finite multihop networks considered, we prove that a Nash equilibrium always exists in the class of centralized strategies. Note that since anonymity, as defined by conditional entropy, is a non-linear function of the probabilities of mixing multiple strategies, the existence of Nash equilibria in classical two-player zero-sum games, where payoff of mixed strategies are weighted sum of pure strategy payoffs, does not directly apply [2]. We demonstrate the applicability of the approach by using a numerical example of a switching network

C. Related Work

Anonymous communication over the Internet is fairly well studied, where many applications have been designed based on the concept of traffic mixes proposed by David Chaum [3]. While mix-based solutions have been used in applications such as anonymous email or browsing, it has been shown that when long streams of packets with latency or buffer constraints are forwarded through mixes, it is possible to correlate incoming and outgoing streams almost perfectly [4]. In wireless networks, an alternative solution to Mixing is the use of cover traffic [5], which ensures that the transmission schedules of all nodes are fixed apriori. While the fixed scheduling strategy provides complete anonymity, it was found to be inefficient [5] and requires synchronization across all nodes.

In this work, we adopt the mathematical framework developed in [6] for omniscient adversaries, where equivocation was used to quantify anonymity of routes, and it was shown that anonymity in network communication requires a reduction in network throughput. The general adversary model considered here necessitates a game-theoretic formulation of the problem. Game theory [7] has been used in a wide range of multi-agent problems from economics to networking. In the context of network security, game-theoretic models have primarily been used to model problems related to distributed intrusion detection [8]–[10]. To the best of our knowledge, ours is the first application of game-theory to hiding traffic flows in the presence of eavesdroppers. The work closest to ours in this regard is that of information concealing games using finite dimensional data, which was studied in [11]. In [11], a two player game is considered where one of the players chooses a subset of available resources to maximize his payoff, while the adversary chooses a subset of resources to hide, so that the payoff is minimized. The authors identify conditions under which Nash equilibria exist and provide approximate techniques to compute the equilibria. While this problem has some conceptual similarities to our strategy of covert relaying, our model utilizes conditional entropy— a non linear function of probabilities of mixing strategies— as the payoff and is thus different from classical mixed strategy models [2].

II. System Model

A. Notation

Let the network be represented by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of directed links. (A, B) is an element of $\mathcal{E}$ iff node B can receive transmissions from node A. A sequence of nodes $P = (V_1, \cdots, V_n) \in \mathcal{V}^*$ is a *valid path* in $\mathcal{G}$ if $(V_i, V_{i+1}) \in \mathcal{E},\ \forall i < n$. The set of all paths is denoted by $\mathcal{P}(\mathcal{G})$.

During any network observation by the adversary, a subset of nodes communicate using a fixed set of paths. This set of paths $\mathbf{S} \in 2^{\mathcal{P}(\mathcal{G})}$ is referred to as a network *session*. The adversary's goal is to use his observation to identify the session. We model $\mathbf{S}$ as an i.i.d. random variable $\mathbf{S} \sim p(\mathbf{S})$. The prior $p(\mathbf{S})$ on sessions is assumed to be available to the adversary. Let $\mathcal{S}$ denote the set of possible sessions.

B. Transmission Scheduling

Each source node transmits a long stream of packets according to an independent transmission schedule. The schedule of node A is denoted by $\mathbf{t}_A = (t_A(1), t_A(2), \cdots)$, where $t_A(i)$ is the time point of transmission of the i^{th} packet by node A. Let λ_A denote the transmission rate of node A:

$$\lambda_A = \lim_{n \to \infty} \frac{n}{t_A(n)}.$$

Transmitter Directed Signaling We consider orthogonal transmitter directed signaling at the physical layer, where each node utilizes a unique orthogonal signaling scheme such that transmissions by multiple nodes are non-interfering. Consequently, the bandwidth constraints correspond to a per-node rate constraint; for every node $A \in \mathcal{V}$, the packet transmission rate λ_A is bounded by a constant C_A. We

assume that the network operates in full duplex mode.

Observable Scheduling The adversary can detect packet transmission times of a subset of nodes, denoted by $\mathbf{N}_e$. Let $\mathbf{t}^e = \{\mathbf{t}_A : A \in \mathbf{N}_e\}$ denote the adversary's complete observation. The adversary's choice is subject to a constraint on the maximum number of monitored nodes, denoted by k_e (also referred to as *power of the adversary*). We model $\mathbf{N}_e$ as a random variable where the random distribution of $\mathbf{N}_e$ is chosen by the adversary to maximize his payoff.

Delay Constraints and Relaying Schedule We impose a *quality of service* requirement on the traffic latency, where every data packet received by a relay must be forwarded within Δ time units or otherwise dropped. The transmission schedule of each node includes the time points of all data packets (some of which could be dropped at subsequent nodes) as well as dummy packets. For every node A, let $\mathbf{t}^r_A$ denote the subset of transmission times by node A that correspond to data packets that are ultimately relayed to the destination. The set of schedules $\{\mathbf{t}^r_A : A \in \mathcal{V}\}$ should satisfy delay constraints at each relay, and for any source node A_s in the session,

$$\lambda^r_{A_s} = \lim_{n\to\infty} \frac{n}{t^r_{A_s}(n)}$$

measures the rate of relayed data packets.

C. Performance Metrics: Anonymity and Throughput

The task of the network designer is to design the set of transmission schedules $\mathbf{t} = \{\mathbf{t}_A : A \in \mathcal{V}\}$ and the subset $\mathbf{t}^r = \{\mathbf{t}^r_A : A \in \mathcal{V}\}$ corresponding to the relayed data packets, such that a desired network throughput is achieved while the adversary obtains minimum information about the session $\mathbf{S}$ by observing $\mathbf{t}_e$. We use $\mathbf{T}, \mathbf{T}^r, \mathbf{T}^e$ as the random variables that denote the transmission schedule, relaying strategy and observable schedule respectively, and model the strategy of the network designer by a probability distribution $q_n(\mathbf{t}, \mathbf{t}^r|\mathbf{s})$.

The task of the adversary is to design the probability distribution $q_e(\mathbf{n}_e)$ of monitored nodes such that maximum information can be obtained by observing $\mathbf{t}^e$.

Anonymity We quantify anonymity using Shannon's equivocation which measures the uncertainty of the sessions given the complete observation of the adversary.

Definition 1: We define the *anonymity* $A(q_n, q_e)$ of a scheduling strategy $q_n(\mathbf{t}, \mathbf{t}^r|\mathbf{s})$ w.r.t adversary strategy $q_e(\mathbf{n}_e)$ as the normalized conditional entropy of the sessions given the adversary observation:

$$A(q_n, q_e) \triangleq \frac{H(\mathbf{S}|\mathbf{T}^e)}{H(\mathbf{S})}. \quad (1)$$

The motivation behind the above definition comes from Fano's inequality which lower bounds the adversary's probability of error by the conditional entropy [12].

Throughput Let the source nodes in session $\mathbf{s}$ be denoted by $A_1, A_2, \cdots, A_{|\mathbf{s}|}$. Then the set of data rates achieved by a relay schedule $\mathbf{t}^r$ in the session $\mathbf{s}$ is given by $\mathcal{L}(\mathbf{s}, \mathbf{t}^r) = (\lambda^r_{A_1}, \cdots, \lambda^r_{A_{|\mathbf{s}|}})$, and the sum-rate in the session is

$$\Lambda(\mathbf{s}, \mathbf{t}^r) = \sum_i \lambda^r_{A_i}.$$

Definition 2: We define the *throughput* $\Upsilon(q_n)$ of a scheduling strategy $q_n(\mathbf{t}, \mathbf{t}^r|\mathbf{S})$ as the average sum-rate of relayed data packets across the sessions

$$\Upsilon(q_n) = \mathbb{E}\left(\Lambda(\mathbf{S}, \mathbf{T}^r)\right) \quad (2)$$

where the expectation is over the joint pdf of $\mathbf{T}, \mathbf{T}^r$ and $\mathbf{S}$.

Anonymity and throughput are essentially two opposing paradigms in the design of the optimal scheduling strategy; transmitting more dummy packets increases anonymity while higher throughput necessitates fewer dummy transmissions. Unlike the omniscient adversary setup, the uncertainty in the identities of the monitored nodes, *i.e.* the randomness in $\mathbf{N}_e$, complicates the design of the optimal covert relaying strategy, as was illustrated in the example in Section I. We therefore formulate this problem as a two-player zero sum game, and try to establish conditions for the existence of Nash equilibria. In the following section, we describe the game-theoretic formulation of the problem.

III. Two Player Game using Covert Relaying Strategy

We pose the problem as a two-player zero sum game, defined by a $3-$tuple $(\mathcal{A}_n, \mathcal{A}_e, \phi)$ where $\mathcal{A}_n$ and $\mathcal{A}_e$ denote the action spaces of the network designer and the adversary respectively, and $\phi : \mathcal{A}_n \times \mathcal{A}_e \mapsto [0, 1]$ is the payoff function for the network designer (the adversary's payoff is $-\phi(\cdot, \cdot)$).

A. Action Spaces

In its most general form, the action space for the network designer would include the set of all probability distributions $q_n(\mathbf{T}, \mathbf{T}^r|\mathbf{S})$ which is a distribution over the space of point processes. In this work, we restrict ourselves the set of *covert relaying strategies* where each relay node belong to one of two categories: *covert* or *visible*.

Covert relay A covert relay B generates an outgoing transmission schedule that is statistically independent of the schedules of all nodes occurring previously in paths that contain B. Due to statistical independence, no adversary can detect the flow of traffic through a covert relay.

Visible relay: A visible relay B transmits every received packet immediately upon arrival thereby ensuring all arriving packets are relayed successfully within the latency constraint. However, the traffic flow through the visible relay is easily detected by an eavesdropper.

Due to latency and bandwidth constraints, maintaining independent schedules would require a covert relay to drop packets or add dummy packets thereby reducing the rate, whereas visible relays can relay every packet without any rate loss.

In a session $\mathbf{s}$, let $\Lambda^c(\mathbf{s},\mathbf{b})$ denote the achievable sum-rate when $\mathbf{b}$ is the set of covert relays. The characterization of the exact rate loss is not necessary for this exposition, and we will treat it as an abstract quantity. We model the set of covert relays by a random variable $\mathbf{B}_n$ with distribution $q_n(\mathbf{b}_n|\mathbf{s})$ and the class of covert relaying strategies is defined by the set of all distributions $\{q_n(\mathbf{b}_n|\mathbf{s})\}$.

For a given strategy $q_n(\mathbf{b}|\mathbf{s})$, the throughput Υ can be expressed as a linear function:

$$\Upsilon(q_n) = \sum_{\mathbf{s}\in\mathcal{S}} p(\mathbf{s}) \sum_{\mathbf{b}\in 2^{\mathcal{V}}} q_n(\mathbf{b}|\mathbf{s})\Lambda^c(\mathbf{s},\mathbf{b}).$$

By restricting ourselves to the class of covert relaying strategies, we define the action spaces for the network designer and the adversary in the game as follows:

$$\mathcal{A}_n = \left\{ \begin{array}{c} \{q_n(\mathbf{s},\mathbf{b}_n) : \mathbf{s}\in\mathcal{S}, \mathbf{b}_n \subset \mathcal{V}\} : \\ \Upsilon(q_n) \geq \gamma \\ q_n(\mathbf{s},\mathbf{b}_n) \geq 0, \forall \mathbf{s}, \mathbf{b}_n \\ \sum_{\mathbf{b}_n} q(\mathbf{s},\mathbf{b}_n) = 1, \forall \mathbf{s} \end{array} \right.$$

$$\mathcal{A}_e = \left\{ \begin{array}{c} \{q_e(\mathbf{n}_e) : \mathbf{n}_e \in \mathcal{V}^{k_e}\} \\ q_e(\mathbf{n}_e) \geq 0, \forall \mathbf{n}_e \\ \sum_{\mathbf{n}_e} q_e(\mathbf{n}_e) = 1 \end{array} \right.$$

The task of the two participants is to design q_n, q_e to maximize their respective payoffs. The key constraint in the action of the network designer is the throughput requirement ($\Upsilon(q_n) \geq \gamma$). The key constraint for the adversary's action is the maximum number of monitored nodes k_e.

B. Payoff

For any pair of actions $(q_n(\mathbf{s},\mathbf{b}_n), q_e(\mathbf{n}_e))$ of the network designer and adversary respectively, the payoff is the anonymity, the definition of which requires a characterization of the adversary's observation in a session.

Equivalent Adversary Observation At any covert relay, an adversary can not correlate the outgoing transmission schedule with that of any node occurring prior in the path. Hence, for every path through the covert relay, the adversary effectively observes two paths, one terminating prior to the covert relay and one commencing at the covert relay. Given the session and set of covert relays, the adversary's observation $\mathbf{O}_e : \mathcal{S} \times 2^{\mathcal{V}} \times 2^{\mathcal{V}} \mapsto 2^{\mathcal{P}}(\mathcal{G})$ is given by

$$\mathbf{O}_e(\mathbf{s},\mathbf{b}_n,\mathbf{n}_e) = \{\mathbf{p}\cap\mathbf{n}_e : \mathbf{p}\in\mathcal{P}(\mathcal{G}) \text{ and}$$

there exists $\mathbf{p}'\in\mathbf{S}$ such that $\mathbf{p}' = \{\mathbf{p}_1, B_1, \mathbf{p}_2, B_2, \mathbf{p}_3\}$ and one of the following statements are true:
1. $\mathbf{p} = \{\mathbf{p}_1\}$, $B_1 \in \mathbf{b}_n$ and $\mathbf{p}_1 \cap \mathbf{b}_n = \phi$.
2. $\mathbf{p} = \{B_1, \mathbf{p}_2\}$, $B_1, B_2 \in \mathbf{b}_n$ and $\mathbf{p}_2 \cap \mathbf{b}_n = \phi$.
3. $\mathbf{p} = \{B_2, \mathbf{p}_3\}$, $B_2 \in \mathbf{b}_n$ and $\mathbf{p}_3 \cap \mathbf{b}_n = \phi$.
4. $\mathbf{p} = \mathbf{p}'$, and $\mathbf{p}' \cap \mathbf{b}_n = \phi$.

$\mathbf{O}_e(\mathbf{s},\mathbf{b}_n,\mathbf{n}_e)$ is the set of routes observed by the adversary when the nodes in $\mathbf{n}_e$ are monitored during a session $\mathbf{s}$ where $\mathbf{b}_n$ is the set of covert relays. Although the adversary observes the transmission timing of all nodes in the network, it is sufficient for him to use $\mathbf{O}_e(\cdot)$ to obtain his best estimate of the network session. This follows from Lemma 2 in [6], where the partial paths were proven to be the sufficient statistic to detect $\mathbf{s}$.

Define $\mathcal{F}_e : \mathcal{P}(\mathcal{G}) \times 2^{\mathcal{V}} \mapsto 2^{\mathcal{S}\times 2^{\mathcal{V}}}$ to be the pre-image:

$$\mathcal{F}_e(\mathbf{p},\mathbf{n}_e) = \{(\mathbf{s},\mathbf{b}) : \mathbf{O}_e(\mathbf{s},\mathbf{b},\mathbf{n}_e) = \mathbf{p}\}.$$

In other words, $\mathcal{F}_e(\mathbf{p},\mathbf{n}_e)$ is the set of possible pairs of session and covert relays given the observation $\mathbf{p},\mathbf{n}_e$.

For a given pair of strategies $(q_n(\mathbf{s},\mathbf{b}_n), q_e(\mathbf{n}_e)) \in \mathcal{A}_n \times \mathcal{A}_e)$, the payoff function $\phi(q_n, q_e)$ is the anonymity which from Definition 1 is given by:

$$\begin{aligned} \phi(q_n,q_e) &= \frac{H(\mathbf{S}|\mathbf{T}_e)}{H(\mathbf{S})} = \frac{H(\mathbf{S}|\mathbf{O}_e(\mathbf{S},\mathbf{B},\mathbf{N}_e),\mathbf{N}_e)}{H(\mathbf{S})} \\ &= -\frac{1}{H(\mathbf{S}} \sum_{\mathbf{n}_e\in 2^{\mathcal{V}}} \sum_{\mathbf{s}\in\mathcal{S},\mathbf{b}_n\in 2^{\mathcal{V}}} q_e(\mathbf{n}_e)p(\mathbf{s}) \times \\ &\quad q_n(\mathbf{s},\mathbf{b}_n) \log q_{ap}(\mathbf{s},\mathbf{O}_e(\mathbf{s},\mathbf{b}_n,\mathbf{n}_e),\mathbf{b}_e) \quad (3) \end{aligned}$$

where

$$q_{ap}(\mathbf{s},\mathbf{p},\mathbf{n}_e) = \frac{q_n(\mathbf{n}_e,\mathbf{s})p(\mathbf{s})}{\sum_{(\mathbf{s}',\mathbf{b}')\in\mathcal{F}_e(\mathbf{p},\mathbf{n}_e)} q_n(\mathbf{s}',\mathbf{b}')p(\mathbf{s}')}$$

is the aposteriori probability that the current session is $\mathbf{s}$ given the adversary observation $(\mathbf{p},\mathbf{n}_e)$.

It is clear that the interests of the network designer and the adversary are exactly the opposite; while the network designer would prefer to make the monitored nodes covert, the adversary would prefer to monitor the nodes that are not covert. We wish to determine if there is an operating point in the pair of action spaces, where neither the network designer nor the adversary has any incentive to change their strategy, in other words, if this game has a Nash equilibrium.

Definition 3: A pair of strategies $(q_n, q_e) \in \mathcal{A}_n \times \mathcal{A}_e$ constitute a *Nash equilibrium* if:

$$\phi(q_n,q_e) = \sup_{q\in\mathcal{A}_n} \phi(q,q_e) = \inf_{q\in\mathcal{A}_e} \phi(q_n,q). \quad (4)$$

Note that, although it has been shown that two player zero sum games, as defined classically [2], always have a Nash equilibrium in the class of mixed strategies, the result does not extend to the game defined here. While the payoff for a mixed strategy in classical two player games is a weighted sum of the mixing probabilities, in our setup, the payoff is a non-linear functional of the mixing probabilities, as given in (3). The existence of a Nash equilibrium in the defined game is shown in the following theorem.

Theorem 1: 1. For the two player zero-sum game defined by the action spaces $\mathcal{A}_n, \mathcal{A}_e$ and payoff function ϕ, there exists a Nash equilibrium.

Proof: Refer to Appendix. □

The equilibrium condition guarantees that at the operating point, the adversary can use no other strategy to decrease

the anonymity. Characterizing the optimal strategy for the adversary is particularly helpful in networks where additional protection can be provided to nodes that are more likely to be monitored. Note that since the action spaces are defined on the probability simplex, any pair of strategies in $\mathcal{A}_e \times \mathcal{A}_n$ corresponds to a random choice of deterministic strategies, and the Nash equilibrium would therefore be a "pure" (albeit random) strategy equilibrium on the defined action spaces.

The omniscient adversary setup is a specific instance of this game, when the adversary has exactly one action: monitor all nodes in the network. The existence and uniqueness of the Nash equilibrium is trivial in that instance and the operating point is given by the rate distortion optimization [6]:

$$\phi(\gamma) = H(\mathbf{S}) - \inf_{\Upsilon(q_n) \leq \gamma} I(\mathbf{S}; \mathbf{O}_e(\mathbf{S}, \mathbf{B}, \mathcal{V})). \quad (5)$$

In general, the equilibrium may not be unique. In such situations, from the network designer's perspective, it would be useful to obtain the equilibrium point with the absolute maximum payoff. Although computing Nash equilibrium strategies is hard since action spaces are continuous, in many networks it is possible to utilize network structure and session models to characterize the optimal throughput-anonymity tradeoffs and Nash equilibrium strategies. In [13], we consider one such class of *parallel relay networks* where we characterize the optimal strategies for the players and also prove that the Nash equilibrium throughput-anonymity tradeoffs are unique.

IV. Numerical Example: Multihop Multiaccess Switching Network

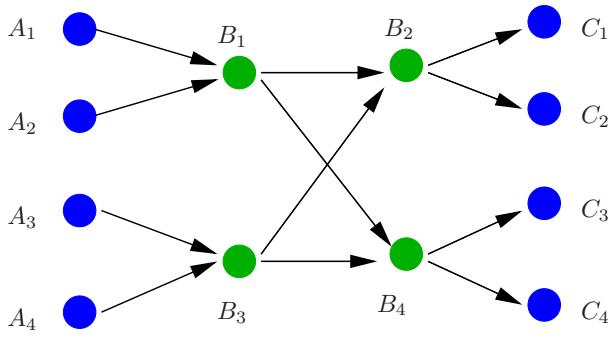

Fig. 2. Switching Network: $\{A_i\}$ transmit to $\{C_i\}$ through relays $\{B_i\}$.

We consider the example of a switching network as shown in Figure 2 which involves multihop routes and multiplexing relays and use numerical methods to study the Nash equilibrium strategies. In any session of the network, each source node A_i picks a unique and distinct destination C_j to transmit packets to. Given a set of source-destination pairs, the set of routes are therefore fixed.

For the network in Figure 2, we assume that all 24 sessions (different source-destination pairings) are equally likely. For the multiaccess relay, the covert relaying strategy used is the priority scheduling algorithm in [14], and the rates are computed assuming all transmission capacities are equal. Figure 3 plots the trade-off between throughput and anonymity for different values of adversarial power k_e.

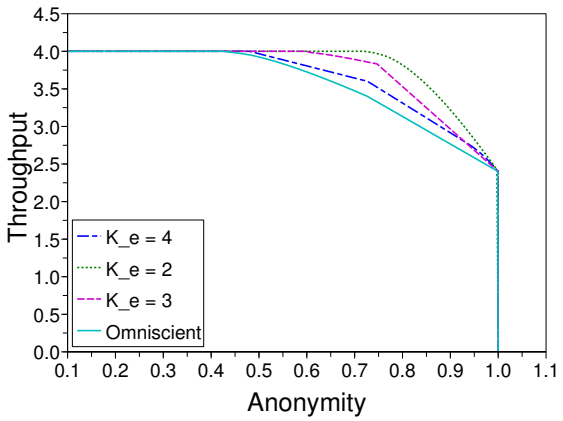

Fig. 3. Throughput-Anonymity Trade-offs for switching network with localized adversary.

In Figure 3, as the adversary's power decreases, the improvement in performance is evident from the figure. For any power of the adversary, maximum anonymity is achievable at the same level of throughput. This is because, as long as the adversary monitors at least two nodes in the network, non-zero information is always revealed unless all relays are covert in all the sessions. Note that the performance of the adversary with power greater than $k_e = 5$ is identical to that of the omniscient adversary. This can be viewed as a fundamental limitation of the topology, wherein it is sufficient for the adversary to monitor at most 5 nodes to obtain maximum information from the network. For every level of adversarial power, there exists a minimum level of anonymity achievable with zero loss in throughput. The reason for this minimum anonymity is the transmission directed signaling which does not reveal the final destination nodes of any transmitted packet.

K_e	Eavesdropper support	Designer Support
2	$\{ \ (B_1, B_3), (B_1, B_4) \ (B_2, B_3), (B_2, B_4) \ \}$	$\{ \ (B_3), (B_4) \ \}$
3	$\{(B_1, B_3, B_4), (B_2, B_3, B_4) \ (A_1, A_3, B_3), (A_2, A_3, B_4)\}$ or $\{(B_1, B_3, B_4), (B_2, B_3, B_4) \ (A_1, A_4, B_4), (A_2, A_4, B_3)\}$	$\{(B_3, B_4), (B_1, B_2) \ (B_1), (B_2)\}$
4	$\{(A_1, A_2, A_3, B_3), (A_1, A_2, A_3, B_4) \ (A_1, A_3, B_1, B_3), (A_2, A_3, B_2, B_4) \ (A_1, A_3, A_4, B_4), (A_1, A_3, A_4, B_3) \ (A_2, A_4, B_1, B_4), (A_1, A_4, B_2, B_3)\}$	$\{(B_3, B_4), (B_1, B_2) \ (B_1), (B_2)\}$

TABLE I
Optimal support set of strategies for network designer and the localized adversary.

Table I summarizes the optimal strategies of the adversaries with different power levels. Note that for $k_e = 3$, there exists at least two possible equilibrium strategies for the adversary (with identical performance). Although we have specified only

one strategy for other values of k_e, the uniqueness of the Nash equilibrium is not guaranteed for the given network.

V. Concluding Remarks

In this work, we considered the problem of providing anonymity to network communication when adversaries monitor an unknown subset of nodes in the network. We presented a game-theoretic formulation and proved the existence of Nash equilibrium. Although we have used specific examples, and assumed knowledge of topology and sessions, a similar approach for random networks with random connections could shed valuable insights into anonymity in mobile ad hoc networks.

VI. Acknowledgements

The author would like to thank Dr. Anand Sarwate for his valuable inputs which lead to the game-theoretic formulation. This work is supported in part by the National Science Foundation under awards CCF-0635070, CCF-0728872 and the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program DAAD19-01-2-0011.

References

[1] T. He and L. Tong, "Detecting Information Flows: Improving Chaff Tolerance by Joint Detection," in *Proc. 2007 Conference on Information Sciences and Systems*, (Baltimore, MD), March 2007.

[2] H. S. Kuhn, *Classics in Game Theory*. Princeton, NJ: Princeton University Press, 1944.

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

[4] 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.

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

[6] P. Venkitasubramaniam, T. He, and L. Tong, "Anonymous networking amidst eavesdroppers," *IEEE Transactions on Information Theory*, vol. 54, pp. 2770–2784, June 2008.

[7] J. F. Nash, "Equilibrium Points in $n-$Person Games," *Proceedings of the National Academy of Sciences*, vol. 36, pp. 48–49, January 1950.

[8] T. Alpcan and T. Basar, "A Game-Theoretic Analysis of Intrusion Detection in Access Control Systems," in *Proc. of 2004 IEEE Conference on Decision and Control.*, (Paradise Island, Bahamas), Dec. 2004.

[9] Y. Liu, C. Comaniciu, and H. Man, "Modeling misbehaviour in adhoc networks: A game-theoretic approach to intrusion detection," *International Journal of Security and Networks*, vol. 1, no. 3-4, pp. 243–254, 2006.

[10] K. Lye and J. M. Wing, ""Game Strategies in Network Security"," *International Journal of Information Security*, vol. 4, pp. 71–86, Feb. 2005.

[11] S. Sarkar, E. Altman, R. El-Azouzi, and Y. Hayel, "Information Concealing Games in Comunication Networks," in *Proc. IEEE INFOCOM*, (Phoenix, AZ), pp. 2119–2127, April 2008.

[12] T. Cover and J. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.

[13] P. Venkitasubramaniam and L. Tong, "A Game-Theoretic Approach to Anonymous Networking." to be submitted to *IEEE Trans. Networking*, 2009.

[14] P. Venkitasubramaniam, T. He, and L. Tong, "Relay Secrecy in Wireless Networks with Eavesdroppers," in *Proc. of 2006 Allerton Conference on Communication, Control and Computing*, (Monticello, IL), Sep. 2006.

[15] J. B. Rosen, "Existence and Uniqueness of Equilibrium Points for Concave $N-$Person Games," *Econometerica*, vol. 33, pp. 520–534, July 1965.

Appendix

A. Proof of Theorem 1

From [15], we know that if in a $2-$player game, the action spaces are closed and convex, and the payoff is continuous and concave in each player's action, then it constitutes a general $2-$player concave game, which is guaranteed to have a Nash equilibrium. We verify these conditions and prove the theorem.

1. **Convexity of action spaces:** The space $\mathcal{A}_e$ is a finite-dimensional simplex, which, by definition is closed, bounded and convex. $\mathcal{A}_n$ is a subset of the simplex with the additional constraint:

$$\Upsilon(q_n) \geq \gamma.$$

Since the constraint is not a strict inequality, the space is closed. $\Upsilon(\cdot)$ is a linear function of q_n. Therefore, for any pair of probability vectors q_n^1, q_n^2

$$\alpha\Upsilon(q_n^1) + (1-\alpha)\Upsilon(q_n^2) = \Upsilon(\alpha q_n^1 + (1-\alpha)q_n^2),$$

which proves the convexity of $\mathcal{A}_n$.

2. Since the payoff is linear in q_e and is an entropy function of q_n, the continuity of the payoff can be easily shown (the details are omitted here).

3. In order to show the concavity of ϕ w.r.t. to q_n, we need to show that for any $q_n^1, q_n^2 \in \mathcal{A}_n$,$q_e \in \mathcal{A}_e$,

$$\alpha\phi(q_n^1, q_e) + (1-\alpha)\phi(q_n^2, q_e) \leq \phi(\alpha q_n^1 + (1-\alpha)q_n^2, q_e).$$

Consider the following modification to the setup, where apart from the topology and set of network sessions, the network designer and the adversary are given access to a common Bernoulli random variable $Z \sim \mathcal{B}(\alpha)$. Consider any $q_n^1, q_n^2 \in \mathcal{A}_n$. The network designer utilizes the following strategy: If the observed variable $Z = 1$, then the distribution q_n^1 is used to make relays covert, and if $Z = 0$, q_n^2 is used. Since Z is observed by the adversary as well, this strategy would amount the anonymity being equal to the conditional entropy $H(\mathbf{S}|\hat{\mathbf{S}}, Z)$.

Now, suppose the Bernoulli variable were only available to the network designer, and he utilizes the same strategy. Since the adversary has no knowledge of Z, his entropy would be $H(\mathbf{S}|\hat{\mathbf{S}})$ where the distribution of covert relays would be the effective distribution:

$$\alpha q_n^1 + (1-\alpha)q_n^2$$

. Since conditioning reduces entropy, $H(\mathbf{S}|\hat{\mathbf{S}}, Z) \leq H(\mathbf{S}|\hat{\mathbf{S}}$, and therefore,

$$\alpha\phi(q_n^1, q_e) + (1-\alpha)\phi(q_n^2, q_e) \leq \phi(\alpha q_n^1 + (1-\alpha)q_n^2, q_e).$$

4. For any q_n, $\phi(q_n, q_e)$ is a linear function of q_e, and therefore,

$$\alpha\phi(q_n, q_e^1) + (1-\alpha)\phi(q_n, q_e^2) = \phi(q_n, \alpha q_e^1 + (1-\alpha)q_e^2),$$

which establishes the required concavity.