Embedding Covert Information Flow

Stefano Marano[†], Vincenzo Matta[†], and Lang Tong[‡]
[†] Department of Information and Electrical Engineering, University of Salerno, Fisciano, SA, Italy
[‡] School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA

***Abstract*— The problem of embedding a covert information flow in independent renewal cover traffic is considered. Such embedding provides maximum anonymity against traffic analysis. The maximum embedding efficiency is characterized, and an accurate approximation is obtained by formulating the problem as a Riemann-Hilbert boundary value problem.**

***Index Terms*—Anonymous networking, traffic analysis, Riemann-Hilbert problem, Fredholm integral equation.**

I. Introduction

Traffic analysis, one of the oldest techniques in surveillance, extracts patterns of communications from traffic measurements. While contents of communications can be protected by cryptographic techniques, patterns of communication and networking can often be inferred from timings of transmissions and the statistics of packet flows. For example, by correlating timings of transmissions at two different nodes, an eavesdropper can draw a conclusion that the two monitored nodes communicate with each other. Similarly, from timing records of a set of nodes, it is possible to conclude that these set of nodes form a route for a particular information flow.

A classical way of hiding communications from traffic analysis is using a certain cover traffic as a carrier for the intended information flow. If all the transmissions are encrypted, an eavesdropper is not able to infer whether a particular transmission is part of an information flow or it is merely one for some dummy packet or a transmission of some multiplexed traffic. If the cover traffic is designed in such a way that transmission epochs at two nodes are statistically independent, then it is impossible for an eavesdropper to make any inference about whether the two monitored nodes are communicating with each other or they are communicating independently with their respective neighbors. Suppose that the two nodes can communicate using a subset of transmission epochs generated from independent transmission schedules, their communication is perfectly anonymous against traffic analysis.

We study in this paper the efficiency of embedding low latency information flow in renewal cover traffic. Consider the case when two nodes A and B use independent transmission schedules as cover traffic. Suppose that the rate of transmissions of both nodes are the same, say λ packets per second. If the information flow has no constraints on packet delays, almost all transmission epochs can be used for relaying packets of the information flow, and the transmission efficiency is 100%. On the other hand, for a low latency flow, if there is a strict delay constraint on the transmission of all the packets, then only some transmission epochs are eligible for relaying packets of the anonymous flow. What, then, is the maximum rate of information flow that can be embedded in a given cover traffic, and what kind of cover traffic is most effective in hiding information flow? To this end, this paper presents some answers and insights when the cover traffic are independent renewal processes.

A. Summary of Results and Contributions

Due to space limitation, this paper contains a summary of results with abbreviated exposition. Technical details can be found in [1].

We consider the problem of embedding of an information flow with strict delay constraint Δ in independent and identical renewal cover traffic at two relaying nodes. In particular, given realizations of independent renewal processes, $S = (s_i)$ at the source and $R = (r_i)$ at the relay, we are intereseted in the optimal embedding strategy that chooses the largest subsets of transmission epochs in S and R to carry the information flow. For the given cover traffic, such an embedding gives the highest efficiency—referred to as *embedding capacity*—of carrying perfectly anonymous information flow.

The problem of optimal embedding, fortunately, has been solved by Blum *et al.* [2] who showed that a simple greedy strategy, referred to as the Bounded Greedy Match (BGM), is sample path optimal. While BGM is easy to implement, the analysis of its performance is nontrivial, and the lack of analytical characterization is a main road block to understanding factors affecting the maximum rate of anonymous information flow.

The main contribution of this paper lies in a simple characterization of the optimal efficiency of BGM for the general renewal cover traffic. In particular, we obtain a sharp approximation of the embedding capacity C^* as

$$C^* \approx \frac{\lambda\Delta}{1 + \frac{2}{\lambda\Delta}\int_0^{\lambda\Delta} m(t)dt} \tag{1}$$

where Δ is the flow delay constraint, λ the rate of cover traffic, and $m(t)$ the renewal function of the underlying renewal processes.

Equation (1) shows the striking fact that it is the renewal function, not the specific distributions of the inter-arrival variable, that plays the crucial role in determining the embedding capacity in renewal cover traffic. Since the renewal function is at the center of renewal theory, properties of $m(t)$ are well

This work is supported in part by the Army Research Office MURI program under award W911NF-08-1-0238.

studied and understood. In many cases of practical interest, the integral involved in (1) can be computed in closed form.

The source of approximation in (1) arises from the omission of higher order terms in the Whittaker-Shannon interpolation formula used in exact expression of C^*. Although additional terms can be included for better accuracy, the simple expression in (1) that connects to a key parameter of the renewal process makes (1) appealing. Furthermore, for Poisson cover traffic, the approximation becomes strictly equal, and the right hand side of (1) matches with C^*.

Fig. 1 shows the accuracy of the approximation in (1) for a number of cover traffic models, which shows excellent match between the analytical approximation given in the right hand side of (1) and C^* obtained through a Monte Carlo evaluation of the embedding capacity. It also shows the preference of one distribution over another. In particular, we see that cover traffic using renewal processes generated by the Gamma distributed interarrival with parameter $\xi = 3$ has a meaningful gain over the Poisson cover traffic, especially at the low latency regime ($\delta < 1$)

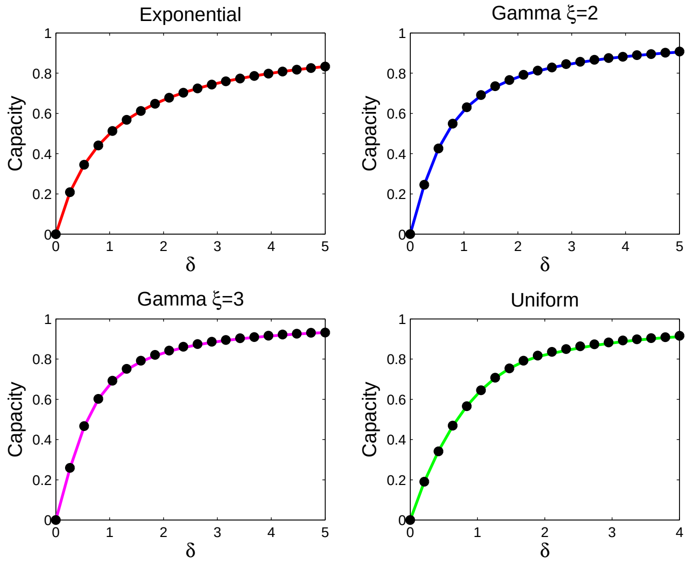

Fig. 1. The embedding capacity for typical cover traffic models. The solid lines are theoretic evaluation from (1) and points are Monte Carlo evaluation of the actual C^*. δ is the delay bound for the flow normalized by the rate of cover traffic.

Another contribution of this work is a new approach based on powerful tools from the Riemann-Hilbert theory. By casting the integral equation that models the dynamics of BGM algorithm as a form of a Riemann-Hilbert problem, we are able to reduce the problem of solving a Fredholm integral equation to one of solving a linear system.

B. Related Work

Traffic analysis has a long storied history, and it played a critical role in modern warfare [3]. The use of traffic analysis in computer and communication networks is well documented [4], [5], [6], [7], [8]. Traffic analysis has shown to expose security weaknesses of SSH [9] and web browsing [10], [6]. Techniques inspired by information theory have also been reported [11], [7], [12].

The idea of using cover traffic to make information flow anonymous is well known. A fundamental concept in anonymous networking is the notion of *mix* introduced by Chaum [13]. The idea is to disguise the presence of information flow by mixing encrypted packets from different users, reordering them, add "dummy" messages, and transmitting them using a randomized schedule. There has been extensive study of the idea of mixing for high latency traffic such as anonymous mailers and proxies [14], [15], [16], [6].

For traffic with low latency requirement—a scenario of focus in this paper, the reported results are limited, although many heuristic practical techniques exist, see *e.g.,* [17], [18], [19], [20], [21], [22] and references therein. It has been shown [12], [23], both experimentally and in theoretical analysis, that simple mixing techniques such as those summarized in [24] are not effective.

The first formal study of the embedding information flow in independent point processes as a way to provide perfect anonymity was presented in [25] where Poisson processes are used as the cover traffic. In that case, the maximum embedding efficiency was shown to be the right hand side of (1). Characterization of anonymous flows in a multi-flow setting was also given [25]. The results presented in this paper can be considered as a generalization of [25], [26]. Our approach builds upon an earlier attempt [27], [28] where the authors obtained an integral equation that characterizes the maximum embedding efficiency. However, there is no attempt in [27], [28] to obtaining analytically tractable solutions.

II. Problem Formulation

We focus in this paper the problem of embedding a directional information flow in independent renewal cover traffic at two nodes. We assume that all transmitted packets are encrypted, thus no protocol information such as source/destination addresses can be inferred. To avoid traffic analysis, it is also common to make packet length (even the distribution of symbols in a packet) identical across all packets. Therefore, we model transmissions at each node as a point process where each point corresponds to the transmission of a packet. Not modeling packet lengths affords us the powerful tools from the theory of point processes with a minor loss of generality. We follow the convention that capital letters denote random variables, and the corresponding lowercase the associate realizations.

We first formally define the notion of (directional) information flow. In this paper, we consider information flows with low latency by imposing a packet-level delay constraint. Specifically, a relay node must forward each packet in the flow within Δ seconds after the packet has been received. Formally, the notion of information flow with bounded delay constraint is defined as follows.

Definition 1: (Information Flow) Two point processes $\mathcal{W} = (w_1, w_2, \dots)$ and $\mathcal{Z} = (z_1, z_2, \dots)$ form an information flow (in the direction $\mathcal{W} \rightarrow \mathcal{Z}$) with delay bound Δ if for every realization $\{w_i\}$ and $\{z_i\}$, there is an one-one mapping

$\{w_i\} \to \{z_i\}$ that satisfies the causal bounded delay constraint $0 \le z_i - w_i \le \Delta$ for all i. ⋄

Given the cover traffic modeled by a pair of point processes defined over $t \in (0, \infty)$, $\mathcal{S} = (S_1, S_2, \ldots)$ at node A and $\mathcal{T} = (T_1, T_2, \ldots)$ at node B, we are interested in embedding an information flow with delay bound Δ in $(\mathcal{S}, \mathcal{T})$ by selecting a subset of points in $(\mathcal{S}, \mathcal{T})$ for the transmissions of packets associated with the flow. Formally, we define an embedding policy as follows with an illustration shown in Fig. 2.

Definition 2: (Embedding Policy) An embedding policy ϵ assigns transmission epochs (subsequences of $\mathcal{S}$ and $\mathcal{T}$) for information flow. ⋄

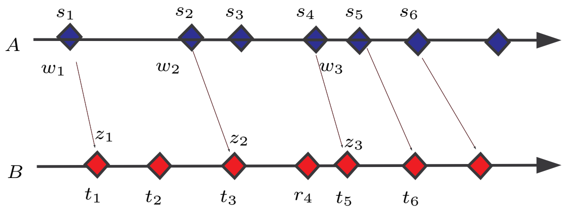

Fig. 2. An illustration of an embedding an information flow $(w_i) \to (z_i)$ in cover traffic $\mathcal{S} = (s_i)$ at node A and $\mathcal{T} = (t_i)$ at node B

Let $\mathcal{E} = \{\epsilon\}$ be the set of admissible embedding strategies. Given $\epsilon \in \mathcal{E}$, the cover traffic $\mathcal{S}$ and $\mathcal{T}$ are decomposed into

$$\mathcal{S} = \mathcal{W}^\epsilon \oplus \mathcal{U}^\epsilon, \quad \mathcal{T} = \mathcal{Z}^\epsilon \oplus \mathcal{V}^\epsilon,$$

where $(\mathcal{W}^\epsilon, \mathcal{Z}^\epsilon)$ forms a valid information flow. Here $\oplus$ is the superposition operator for point processes: $\{c_i\}$= $\{a_i\} \oplus \{b_i\}$ means that $\{c_i\} = \{a_i\} \cup \{b_i\}$ with $c_1 \le c_2 \le \ldots$.

Each particular embedding has a certain capability of secretly hosting information flows. This is quantified as follows.

Definition 3: (Embedding Efficiency) The efficiency of a given embedding $\epsilon \in \mathcal{E}$ is

$$\eta(\epsilon) = \lim_{t \to \infty} \frac{2\, N_{\mathcal{W}^\epsilon}(t)}{N_{\mathcal{S}}(t) + N_{\mathcal{T}}(t)},$$

where $N_{\mathcal{W}^\epsilon}(t)$ is the counting process for the embedded information process $\mathcal{W}^\epsilon$, and $N_{\mathcal{S}}(t)$, $N_{\mathcal{T}}(t)$ are similarly defined. We assume the limit exists almost surely. ⋄

It is obvious that, given a pair of cover traffic with finite means, there is a maximum fraction of points in the cover processes can be used for carrying the information flow. We are interested in the highest efficiency, that we call the *embedding capacity*:

Definition 4: (Embedding Capacity) $C^* = \sup_{\epsilon \in \mathcal{E}} \eta(\epsilon)$. ⋄

Note that C^* is a function of Δ and the point process statistics.

We shall focus on the case that $\mathcal{S}$ and $\mathcal{T}$ are independent renewal processes, with interarrival random variables X and Y, respectively. It is assumed that X and Y are identically distributed, and the rate of the processes is finite and nonzero, $0 < \lambda = 1/\mathbb{E}[X] = 1/\mathbb{E}[Y] < \infty$. When the second moment is finite, we define the dispersion index

$$\gamma = \lambda^2\, \mathrm{VAR}[X] = \lambda^2\, \mathrm{VAR}[Y] < \infty.$$

The efficiency of any Δ-matching rule between point processes with rate λ is only function of the product $\lambda\Delta$. Indeed, doubling the arrival rate "speeds up" the system so that the sample paths can be redrawn on a time axis scaled by a factor 2, and halving Δ leaves unchanged the number of matches. We accordingly introduce the following further notation.

- The *normalized delay* is $\delta = \lambda\Delta$.
- The dependence upon the normalized delay is made explicit by denoting the capacity as $C^*(\delta)$.
- Assuming that the interarrival random variables X and Y admit Lebesgue density, we introduce the density of the normalized (unit-rate) variables $\widetilde{X} = \lambda X$ and $\widetilde{Y} = \lambda Y$, that is denoted by $k(t)$:

$$\int k(t)dt = \int t\, k(t)dt = 1.$$

- The renewal function of the unit-rate processes is denoted by $m(t) = \mathbb{E}[N(t)]$, where $N(t)$ is the number of arrivals in $(0, t)$ of the normalized processes.

III. Characterization of Embedding Capacity

A. Optimal Embedding Policy

As a first step toward capacity evaluation, we need to know whether or not an optimal policy exists that maximizes the number of matched packets for a given kind of cover traffic schedules, thus achieving the embedding capacity.

An optimal algorithm has been found in the literature, which does exactly what we are interested in: the Bounded Greedy Match (BGM) [2]. It is a simple algorithm that classifies the events of two arbitrary point processes as matched and unmatched, where points are sequentially marked as matched if they comply with the causal delay constraint Δ, and as unmatched otherwise. The BGM algorithm works as follows: Two point processes are given; all the points as initially "unmatched." The BGM consists in repeating the following two steps.

1) Consider the first (in the direction of increasing time) unmatched point occurring at the first process, say $p^{(1)}$.
2) The first unmatched point on the second process at distance not larger than Δ from $p^{(1)}$, if any, is denoted by $p^{(2)}$; if such point exists mark both $p^{(1)}$ and $p^{(2)}$ as "matched."

Matched and unmatched points are also referred to as "flow" and "chaff", respectively. The BGM is optimal, in the sense that, given two realizations of arbitrary point processes and arbitrary value of the constraint Δ, the BGM algorithm results in the minimum number of chaff points [25], [26], [2].

B. Main Results

The main contribution of this paper is contained in the following set of Theorems. Theorem 1 gives an expression for $C^*(\delta)$. Theorem 2 provides the analytical approximation of $C^*(\delta)$ as shown in (1). The Corollary 3 shows the scaling behavior of the embedding for large δ.

Theorem 1: (Exact value of $C^*(\delta)$) Under the assumption of finite second moment for the interarrivals, the embedding capacity of two independent and identically distributed renewal processes, under normalized delay constraint δ, is

$$C^*(\delta) = \frac{2\Omega(0)}{1+\Omega(0)}, \qquad (2)$$

where $\Omega(f)$ is the solution of

$$\Omega(f) + 2\int \Omega(\nu)\Re\left\{\frac{K(\nu)}{1-K(\nu)}\right\}\delta\mathrm{sinc}[\delta(f-\nu)]d\nu = \delta\mathrm{sinc}(\delta f)\,\frac{1-\Omega(0)}{2}, \qquad (3)$$

and $K(f)$ is the Fourier transform of the normalized density $k(t)$. ◇

The proof of Theorem 1 involves several steps. First, it is necessary to obtain an analytical model for the optimal embedding algorithm BGM. Following [26], an uncountable state Markov chain model can be constructed, and the stationary distribution u of that Markov chain is related to $C^*(\delta)$ by

$$C^*(\delta) = \frac{2\int_{-\delta/2}^{\delta/2} u(t)dt}{1+\int_{-\delta/2}^{\delta/2} u(t)dt}. \qquad (4)$$

The stationary distribution u is shown in [27] to satisfy a homogeneous Fredholm integral equation of the second kind. Next, we draw a connection of the solution of the specific integral equation that determines u with the celebrated Riemann-Hilbert problem [29]. Finally, using the approaches of Jones [30], see also [31], and an application of the analytic continuation theorem [32], we obtain (3).

In the theory of integral equations it is common practice to transform one equation into another, which is more amenable to exact or approximate solution, and this is just what we make of eq. (3). More important, starting from (3), in the following theorem we provide a fully analytical approximation of the embedding capacity.

Theorem 2: (Approximation of $C^*(\delta)$) Under the assumption of finite second moment for the interarrivals, the embedding capacity of two independent and identically distributed renewal processes, with normalized renewal function $m(t)$ and normalized delay constraint δ, can be approximated as

$$C^*(\delta) \approx C(\delta) = \frac{\delta}{1+\dfrac{2}{\delta}\displaystyle\int_0^\delta m(t)dt}. \qquad (5)$$

◇

The relevance of the above result stems from the fact that, for the typical distributions encountered in many applications, the accuracy of the fully analytical approximation (5) seems to be excellent irrespective of the range of δ and of the distribution *heavyness*. Theorem 2 provides us with an accurate yet mathematically tractable expression for the embedding capacity under arbitrary renewal traffic.

We emphasize that the characterization (5) relates the sought capacity to the renewal function of the underlying process, averaged over an interval δ. This highlights the role of the renewal function $m(t)$, and reveals that its average $\frac{1}{\delta}\int_0^\delta m(t)dt$ is the key quantity in determining $C(\delta)$. Thus, different traffic models can be classified with respect to their embedding capabilities just in terms of that average. Expression (5) also suggests numerical approaches to the capacity computation. For instance, an estimate of $C(\delta)$ can be obtained even without knowing the underlying distribution by simply counting the number of arrivals within a time interval of δ.

We now state a corollary characterizing the asymptotic behavior of the capacity in the limit of large δ. From a known property of the renewal function, $m(t) - t \to (\gamma-1)/2$ in the limit of $t\to\infty$. Plugging that expression in eq. (5) would give $1 - C(\delta) \sim \gamma/\delta$. Indeed, we have the following result.

Corollary 1: (Scaling law for $C(\delta)$) Under the assumption of finite second moment for the interarrivals, $\lim_{\delta\to\infty}[1 - C(\delta)]\delta = \gamma$, i.e., the embedding capacity in Theorem 3 scales as

$$1 - C(\delta) \sim \gamma/\delta.$$

◇

The corollary reveals that, for large δ, the key quantity in determining the capacity is the dispersion index. The ability of masking information flows in independent realizations only depends on the value of the global parameter γ, and different traffic models sharing the same dispersion behave similarly for large normalized delays (namely $\Delta \gg 1/\lambda$).

Finally, to improve on the approximation in Theorem 2, we provide the following theorem that expresses the embedding capacity as the solution of a simple linear system. Consider, for any integer $N \geq 1$, the following system

$$\sum_{k=-N}^{N} A_{hk}\,\Omega(k/\delta) = \frac{\delta}{2} I_h,$$

where $I_h = 1$ for $h = 0$, and $I_h = 0$ otherwise. The analytical expressions of the entries A_{hk}, defining a $2N+1$ by $2N+1$ matrix $\boldsymbol{A}$, are given by

$$A_{00} = 1 - \frac{\delta}{2} + \frac{2}{\delta}\int_0^\delta m(t)dt, \qquad (6)$$

$$A_{kk} = 1 + \frac{2}{\delta}\int_0^\delta m(t)[\cos(2\pi kt/\delta) + 2\pi k\,(1-t/\delta)\,\sin(2\pi kt/\delta)]\,dt, \qquad k \neq 0, \qquad (7)$$

$$A_{0k} = (-1)^k\,\frac{2}{\delta}\int_0^\delta m(t)\,\cos(2\pi kt/\delta)\,dt, \qquad k \neq 0, \qquad (8)$$

$$A_{hk} = \frac{(-1)^{h-k}}{(h-k)}\left[h(-1)^h A_{0h} - k(-1)^k A_{0k}\right], \quad h \neq k. \qquad (9)$$

Theorem 3: (Linear system approximation of $C^*(\delta)$) Under the assumption of finite second moment for the interarrivals, let $C^*(\delta) = \frac{2\Omega(0)}{1+\Omega(0)}$ as in Theorem 2. Then, assuming that $\boldsymbol{A}$ is invertible, $\Omega(0)$ can be approximated as $\delta/2$ times the $(0,0)$-entry of matrix $\boldsymbol{A}^{-1}$, namely $\Omega(0) = \frac{\delta}{2}\{\boldsymbol{A}^{-1}\}_{00}$. In particular, specializing for $N = 1$, the capacity becomes

$$C^*(\delta) \approx \frac{\delta}{1+\dfrac{2}{\delta}\displaystyle\int_0^\delta m(t)dt + 2\,\dfrac{A_{01}^2}{A_{01}-A_{11}}}. \qquad (10)$$

◇

First, note that, in the refined approximation corresponding to $N = 1$, a correction term $2\,\frac{A_{01}^2}{A_{01}-A_{11}}$ appears, with respect to the approximation $C(\delta)$ in eq. (5), which only uses A_{00}. Second, we note that $\boldsymbol{A}$ is very structured and its degrees of freedom grow only linearly with N; in fact, $\boldsymbol{A}$ is completely specified by assigning one row and the main diagonal. This structure is very convenient for numerical tractability. Finally, it is expected that the solution becomes more and more accurate as the system size N increases. In the section devoted to numerical experiments, we show that the *zero-order* approximation $C(\delta)$ is well satisfying in many cases of interest. Even when this is not strictly true, a first-order correction (10) offers very good results.

IV. Conclusion

We consider the problem of matching two independent and identically distributed renewal processes, according to a bounded delay criterion, with applications to communication network scenarios. We introduce the concept of *embedding capacity*, and provide fully analytical tools and approximations to evaluate it, relying upon the Riemann-Hilbert theory. An exact evaluation of the capacity is reduced to a manageable integral equation, that can be solved to any degree of approximation by solving a structured linear system. The main finding, however, is a simple approximated formula of the embedding capacity that involves the renewal function of the underlying processes. The approximation is excellent for virtually all the cases of practical interest that we have investigated, part of which are reported in the paper. Even when this is not strictly true, we provide closed-form solutions for first-order correction.

The abstract concept of matching between point processes arises in a very large number of contexts, and we feel that our findings can represent a contribution to these fields. To broaden further the horizon of potential applications, refinements and improvements of the approach can be considered. These include: the case of different renewal processes at the two nodes, the extension to multi-hop flows, and the case of multiple input/multiple output relays, see [25], [26].

References

[1] S. Marano, V. Matta, and L. Tong, "Embedding anonymous information flow in renewal cover traffic." in prepration.

[2] 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.

[3] N. West, *The SIGINT Secrets: The Signal Intelligence War: 1900 to Today*. New York: William Morrow, 1988.

[4] V. L. Voydock and S. T. Kent, "Security mechanisms in high-level network protocols," *ACM Computing Surveys*, vol. 15, pp. 135–171, 1983.

[5] J.-F. Raymond, "Traffic analysis: Protocols, attacks, design issues and open problems," in *Designing Privacy Enhancing Technologies: Proceedings of International Workshop on Design Issues in Anonymity and Unobservability* (H. Federrath, ed.), vol. 2009 of *LNCS*, pp. 10–29, Springer-Verlag, 2001.

[6] 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.

[7] X. Fu, B. Graham, R. Bettati, and W. Zhao, "On countermeasures to traffic analysis attacks," in *Information Assurance Workshop, 2003. IEEE Systems, Man and Cybernetics Society*, pp. 188–195, 18-23 June 2003.

[8] N. Matthewson and R. Dingledine, "Practical traffic analysis: Extending and resisting statistical disclosure," in *Privacy Enhancing Technologies: 4th International Workshop*, May 2004.

[9] D. X. Song, D. Wagner, and X. Tian, "Timing Analysis of Keystrokes and Timing Attacks on SSH," in *Proc. 10th USENIX Security Symposium*, (Washington, DC), August 2001.

[10] E. W. Felten and M. A. Schneider, "Timing attacks on web privacy," in *ACM Conference on Computer and Communications Security*, pp. 25–32, 2000.

[11] W. Lee and D. Xiang, "Information-theoretic measures for anomaly detection," in *Proceedings of the 2001 IEEE Symposium on Security and Privacy*, vol. 00, (Los Alamitos, CA, USA), p. 0130, IEEE Computer Society, 2001.

[12] Y. Zhu, X. Fu, R. Bettati, and W. Zhao, "Anonymity analysis of mix networks against flow-correlation attacks," in *Proceedings of IEEE Global Communications Conference*, November 2005.

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

[14] 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.

[15] 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.

[16] 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.

[17] O. Berthold, H. Federrath, and S. Kopsell, "Web MIXes: A system for anonymous and unobservable Internet access," in *Proceedings of the Workshop on Design Issues in Anonymity and Unobservability, Lecture Notes in Computer Science*, vol. 2009, (Berkeley, CA), pp. 115–129, July 2000.

[18] M. G. Reed, P. F. Syverson, and D. M. Goldschlag, "Anonymous connections and onion routing," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 482–494, May 1998.

[19] M. Freedman and R. Morris, "Tarzan: A Peer-to-Peer Anonymizing Network Layer," in *Proc. ACM Conference on Computer and Communications Security*, Nov. 2002.

[20] B. Levine, M. Reiter, C. Wang, and M. Wright, "Timing attacks in low-latency mix systems," in *Financial Cryptography: Lecture Notes in Computer Science*, vol. 3110, pp. 251–265, Springer, 2004.

[21] G. Danezis, "The traffic analysis of continuous-time mixes," in *Privacy Enhancing Technologies: Lecture Notes in Computer Science*, vol. 3424, pp. 35–50, Springer, 2005.

[22] V. Shmatikov and M. Wang, "Timing analysis in low-latency mix networks: attacks and defenses," in *Computer SecurityESORICS 2006: Lecture Notes in Computer Science*, vol. 4185, pp. 18–33, Springer, 2006.

[23] 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.

[24] A. Serjantov, R. Dingledine, and P. Syverson, "From a trickle to a flood: Active attacks on several MIX types," in *Proceedings of the Fifth International Workshop on Information Hiding (IH'02), Lecture Notes in Computer Science*, vol. 2578, (Noordwijkerhout, The Netherlands), pp. 36–52, October 2002.

[25] P. Venkitasubramaniam, T. He, and L. Tong, "Anonymous Networking Amidst Eavesdroppers," *IEEE Transactions on Information Theory*, vol. 54, pp. 2770–2784, June 2008.

[26] T. He and L. Tong, "Detection of Information Flows," *IEEE Transactions on Information Theory*, vol. 54, pp. 4925–4945, Nov. 2008.

[27] T. He, A. Agaskar, and L. Tong, "On security-aware transmission scheduling," in *Proc. 2008 IEEE Intl. Conference on Acoustics, Speech and Signal Processing*, pp. 1681–1684, March 31-April 4 2008.

[28] T. He, L. Tong, and A. Swami, "Maximum throughput of clandestine relay," in *Proc. 47th Ann. Allerton Conf. on Comm., Contr., and Compt.*, (Allerton, IL), pp. 1082–1089, Oct 2009.

[29] N. I. Muskhelishvili, *ingular Integral Equations: Boundary Problems of Function Theory and their Application to Mathematical Physics*. Dover Publications, 2008.

[30] D. S. Jones, "Diffraction by a wave-guide of finite length," *Proc Camb Phil Soc*, vol. 48, no. 1, pp. 118–134, 1952.

[31] B. Noble, *Methods Based on the Wiener-Hopf Technique for the Solution of Partial Differential Equations*. New York, NY: AMS Chelsea Publishing, 1988.

[32] W. Rudin, *Real and Complex Analysis*. McGraw-Hill Book Company, 1986.