The Chances of Successful Attacks against Continuous-time Mixes

Gergely Tóth and Zoltán Hornák

Budapest University of Technology and Economics
Department of Measurement and Information Systems
H-1117 Budapest, XI., Magyar tudósok krt. 2.
{tgm,hornak}@mit.bme.hu

Abstract. Continuous-time mixes represent a relatively new field in anonymity services. Their simplicity and probabilistic approach suggest promising results. In this paper their two most recent representatives, the SG Mix and the PROB Channel will be analysed. The aim of this paper is twofold: first it will be shown via analytical means that considering a locally back-tracing observer and the source-hiding property as anonymity measure, the PROB Channel outperforms the SG Mix for MIN/MAX senders in practically relevant scenarios. Then results of simulations will confirm the theoretical arguments.

Keywords: anonymity measurement, SG Mix, PROB Channel, comparison, simulation

1 Introduction

The field of anonymous electronic communication is rapidly evolving. Different types of anonymity systems are trying to meet the never-ending demands of new IT applications, such as e-voting, e-payment, or browsing and e-mails.

We will introduce the class of continuous-time mixes for anonymous communication and will analyse their two most recent representatives: the SG Mix [1] and the PROB Channel [2]. The goal of this paper is to compare these two systems according to their performance in hiding the relationship between sent and delivered messages: the key task in anonymous communication.

For the analysis two methods will be used: first, by *analytical means*, we will derive the source-hiding property for both systems in order to determine the theoretically maximal chance of being successfully compromised. Next, results of *simulations* will be discussed that describe the practical resistance of the systems against a locally back-tracing passive adversary introduced in [2].

Both from the analytical calculations and from the simulations an interesting result emerges: although the SG Mix was proven to be optimal in [3] for Poisson-distributed messages and an entropy-based anonymity metric[1], the

[1] cf. we used MIN/MAX senders (see Section 4.1) rather than Poisson-distributed messages, the source-hiding property for measuring anonymity rather than entropy and the average ratio of successfully matched messages for measuring the performance of the two systems.

PROB Channel outperformed it both considering the source-hiding property and the simulation results.

Organization of the paper The paper is organized as follows: Section 2 will introduce the basic principles and notions of anonymous communication. Then, Section 3 will present the two continuous-time mixes, the SG Mix and the PROB Channel. Afterwards, in Section 4 we will derive via analytical means the guaranteed level of anonymity, i.e. the source-hiding property of both systems. Next, Section 5 presents the results of simulations in practical scenarios. Finally, conclusions and references close the paper.

2 Overview of Anonymous Communication

In this paper techniques for anonymous communication are evaluated. In order to establish a common understanding, first the model of an anonymous message transmission system (AMTS) will be introduced. The section continues with the description of the foundations for measuring anonymity, the basis for comparing anonymity systems. Finally, we will describe the locally back-tracing passive adversary, who will be referred to in the rest of the paper.

2.1 Abstract Model of an Anonymous Message Transmission System

For the purpose of anonymous electronic communication, several anonymous message transmission systems have been proposed. Their structures and modes of operation differ in various aspects, but some common properties are true for most of them. This common basic framework will be defined in the following.

The goal of an AMTS is to deliver *messages* from *senders* to *recipients* so that it becomes algorithmically hard for an *adversary* to link these messages to senders or recipients. Let us look at the formal model:

- Senders ($s_i \in S$) send messages ($\alpha_j \in \varepsilon_S$) at times $t_S(\alpha_j)$ through the AMTS. These messages are fixed-size and are encrypted for the AMTS.
- Messages are independent from each other, i.e. we do not consider message streams. It is a known issue and subject to further research that the systems analysed in this paper will not be able to provide the calculated level of anonymity if the attacker can link several messages together.
- The AMTS receives the messages and performs cryptographic operations on them to obtain a different representation. In order to further confuse the adversary, the AMTS delays and reorders messages. How messages are actually represented is irrelevant for the purposes of this paper (for example, the packet format MINX has recently been proposed in [4]). The main assumption is that the adversary is not able to break the cryptographic functions.
- After the delay, the AMTS delivers the re-encoded messages ($\beta_k \in \varepsilon_R$) to the recipients ($r_l \in R$) at times $t_R(\beta_k)$. These delivered messages also have a common, fixed size, and are encrypted for the recipient.

– The adversary's aim is to either match the delivered messages to senders ($\beta_k \rightarrow s_i = S(\beta_k)$), or the sent messages to recipients ($\alpha_j \rightarrow r_l = R(\alpha_j)$). In order to do this, adversaries may eavesdrop on communication channels and see when messages are being sent or delivered (*passive* adversary), or even influence the network traffic by delaying messages or creating new ones (*active* adversary). However, we assume now that they cannot break the encryption schemes used and thus they cannot use the contents of messages for correlation purposes.

2.2 Measuring Anonymity

Several constructions have already been proposed in the anonymity field, and newer systems are continuously being built as well. In order to objectively compare them with each other, we need to measure the level of anonymity they provide. For this reason different metrics have been proposed:

– In [5] the level of anonymity was defined as the *size of the anonymity set*[2].
– Both in [7] and in [8] *entropy-based* measures were proposed already taking into account that different senders could have sent the messages with different probabilities.

However, it was shown in [9] that neither of these approaches measure anonymity from the user's perspective perfectly. Users prefer the *local* approach, where compromizing some messages with a probability greater than anticipated is already a successful attack. On the other hand, the *global* approach of these entropy-based measures aims to quantify the amount of information that is needed to *unambiguously* identify the subject corresponding to the respective message. In [9] it was also shown that non-desirable systems can exist for both metrics for an arbitrarily high (simple/normalised) entropy under extreme conditions.

Due to the shortcomings of the entropy-based anonymity metrics, in [2] the *source-hiding property* was introduced for sender anonymity. This new approach defines the level of anonymity as the maximal probability with which the adversary may back-trace messages to their senders (1).

Definition 1. *An AMTS is* source-hiding with parameter Θ *if the adversary cannot assign a sender to a delivered message with a probability greater than* Θ*:*

$$\forall_{\beta_k} \forall_{s_l} \left(P_{\beta_k, s_l} \leq \Theta \right) \tag{1}$$

In this paper we will use the source-hiding property for analytically measuring the anonymity provided by the different continuous-time mixes.

[2] The *anonymity set* consists of the potential subjects that might have performed the particular action of interest [6].

2.3 Locally Back-tracing Passive Adversary

For the previously introduced probability-based anonymity metrics, the big question is how an adversary might associate messages with senders, and thus, how the respective probabilities can be calculated.

In anonymity research several types of adversaries have been described. Naturally, the level of anonymity achieved greatly depends on the adversary assumed. In [2], for the PROB Channel (and further generalised for any continuous-time mix), the concept of the locally back-tracing passive adversary was introduced. We will assume him as the attacker of the systems analysed.

In order to simplify further equations, first two sets need to be defined. With μ_{β_k} (2) the set of sent messages is meant that might have left the AMTS as β_k (i.e. they were sent within the appropriate timeframe[3], whereas η_{β_k,s_l} denotes the subset of μ_{β_k}, which was sent by a specific sender s_l (3).

$$\mu_{\beta_k} = \{\alpha_j | \, (t_R(\beta_k) - \delta_{\max}) < t_S(\alpha_j) < (t_R(\beta_k) - \delta_{\min})\} \tag{2}$$

$$\eta_{\beta_k,s_l} = \{\alpha_j | (\alpha_j \in \mu_{\beta_k}) \wedge (S(\alpha_j) = s_l)\} \tag{3}$$

When this passive adversary performs local back-tracing, he calculates the delivered message → sender matching for each delivered message independently. The following equation (4) gives this guess, which is the probability of a certain delivered message β_k being sent by a certain sender s_l:

$$P_{\beta_k,s_l} = \frac{\sum_{\alpha_j \in \eta_{\beta_k,s_l}} f\left(t_R(\beta_k) - t_S(\alpha_j)\right)}{\sum_{\alpha_j \in \mu_{\beta_k}} f\left(t_R(\beta_k) - t_S(\alpha_j)\right)} \tag{4}$$

Of course the attacker chooses s_i as the sender for β_k where $P_{\beta_k,s_i} = \max_{s_l \in S} P_{\beta_k,s_l}$.

3 Continuous-time Mixes

Continuous-time mixes represent a relatively new group of anonymous message transmission systems. The aim of this paper is to compare the SG Mix and the PROB Channel. Both use probability variables (with different density functions) for calculating the delays of messages. This simple solution is quite effective in practice, though different strategies result in different levels of anonymity.

3.1 Model of the Continuous-time Mixes

There are several ways in which an AMTS can provide anonymity for messages. Chaumian MIXes [10] carry out batching for this purpose: they buffer incoming

[3] This notion was originally devised for real-time mixes, where $\delta_{\max}$ means the guaranteed maximal and $\delta_{\min}$ the minimal delay. Since the delay is limited there, the adversary can easily identify the relevant messages. For non-real-time systems approximations need to be considered, which will be shown later in this paper.

messages until a certain condition has been reached (e.g. a given amount of time has elapsed, a given number of messages has arrived etc.), then shuffle the messages randomly and deliver them afterwards.

Continuous-time mixes, on the other hand, follow a different approach:

- in this scenario every message is processed individually, and its delivery does not depend on the arrival of other messages;
- the delay in the AMTS is a probability variable δ with a given density function $f(\delta)$ (where $\int_0^\infty f(\delta)d\delta = 1$);
- message forwarding in the AMTS happens in the following way: after the arrival of a message at the AMTS, it waits the time specified by the delay (according to the probability variable), and then delivers the message to the recipient.

From the mode of operation it is clear that the main characteristic property of continuous-time mixes is the delay strategy, which is the density function $f(\delta)$ of the delay probability variable.

3.2 Delay Strategies

Recently two continuous-time mixes have been introduced in scientific papers, the Stop-And-Go Mix (the SG Mix) and the PROB Channel. Their description is due now, followed by an analysis of the level of anonymity they provide.

The SG Mix The SG Mix proposed by Kesdogan *et al.* [1] was created with the aim of defining a *probabilistically secure* anonymity system. Under the assumption of Poisson-distributed incoming messages, the SG Mix was proven to achieve this goal. The basic idea is that each sender attaches three parameters to every message sent:

- $TS^{\min}$ and $TS^{\max}$ indicate for the SG Mix that the message is supposed to arrive there between these times. If the message arrives earlier or later, it should be discarded;
- the third parameter attached to a message is the actual requested delay T_i, which is chosen from an exponential distribution with parameter μ. Thus the delay characteristic of the SG Mix is:

$$f(\delta) = \mu \cdot e^{-\mu \cdot \delta} \tag{5}$$

From this it follows that the mean delay of such a system is $\frac{1}{\mu}$.

When a message arrives at the SG Mix, it first checks the parameters $TS^{\min}$ and $TS^{\max}$. If the message has arrived outside the specified time interval, then the AMTS will discard it, since most likely it was delayed by an adversary. If the message looks OK, then the SG Mix will delay it as requested and afterwards deliver it to its recipient.

The PROB Channel The PROB Channel introduced in [2] adopts a different approach:

- in this setting the delay for the messages is specified by the AMTS rather than by the senders;
- in order to achieve real-time properties, the density function is limited by a maximal delay $\delta_{\max}^{\mathrm{PROB}}$, and to indicate that the channel needs a certain amount of time for processing, a minimal delay is also present ($\delta_{\min}^{\mathrm{PROB}}$);
- finally, under these circumstances the uniform distribution was suggested:

$$f(\delta) = \begin{cases} \frac{1}{\delta_{\max}^{\mathrm{PROB}} - \delta_{\min}^{\mathrm{PROB}}} & \text{for } \delta_{\min}^{\mathrm{PROB}} \leq \delta \leq \delta_{\max}^{\mathrm{PROB}}, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

4 Analytical Evaluation of Continuous-time Mixes

In this section the anonymity of the two previously introduced continuous-time mixes will be viewed from the analytical perspective. Based on a locally back-tracing passive adversary, equations for the source-hiding properties will be derived. In the following Section simulation results will demonstrate the strength of the two algorithms.

4.1 Traffic Model

When anonymous message transmission systems are analysed, it has to be considered, how messages arrive at the AMTS. This section describes our approach to the traffic model for the continuous-time mixes to be analysed.

Poisson-distributed Messages In most anonymity related scientific papers, when analysing traffic or measuring the provided level of anonymity it is assumed that messages arrive according to the Poisson distribution. However, in this paper we will analyse a different scenario, that of the MIN/MAX senders. The reason for this is twofold:

- In [11] Díaz *et al.* show that the traffic seen in real-world deployed anonymity systems does not resemble the Poisson distribution. The authors could not characterise the actual traffic as any of the well known distributions.
- The second reason is that in several cases together with the Poisson distribution it is assumed that each message originates from a different sender, which is not realistic in practical systems.

MIN/MAX Senders For the reasons above we do not assume messages arriving at the channel in a Poisson distribution, but we will rather consider MIN/MAX senders: the solution from [2] requires that each sender sends at least one message in each $\tau_{\max}$ time interval (minimal message-sending frequency),

and that no sender sends more than one message in any $\tau_{\min}$ time (maximal message sending frequency). If a sender does not have real messages to send, he should send dummy messages to randomly chosen recipients.

In [12] we concluded that without the MAX rule no anonymity can be given by a real-time channel. However, since the PROB Channel is real-time, we have to require the MAX rule. Therefore, in order to be able to compare the two systems we will use the same environment for both of them.

If senders conform to the MIN/MAX rules, then for (4) a guaranteed upper limit can be given, and thus a reasonable source-hiding property can be achieved. Equation (7) takes the worst case into account, when only one sender sends at the maximal frequency (thus with $\tau_{\min}$), and these messages get delivered at the most probable time intervals according to $f(\delta)$, whereas all the other senders send with minimal frequency (thus with $\tau_{\max}$), and their messages get delivered at the least probable intervals. (This assumption is only valid if $\tau_{\max} \leq \delta_{\max}$).

$$\Theta = \frac{\sum_{i=1}^{\Delta_{\min}} \max_{\delta_{\min}+(i-1)\cdot\tau_{\min} \leq q \leq \delta_{\min}+i\cdot\tau_{\min}} f(q)}{|S| \cdot \sum_{i=1}^{\Delta_{\max}} \min_{\delta_{\min}+(i-1)\cdot\tau_{\max} \leq q \leq \delta_{\min}+i\cdot\tau_{\max}} f(q)} \tag{7}$$

where $\Delta_{\max} = \lfloor \frac{\delta_{\max}-\delta_{\min}}{\tau_{\max}} \rfloor$ and $\Delta_{\min} = \lceil \frac{\delta_{\max}-\delta_{\min}}{\tau_{\min}} \rceil$.

4.2 Anonymity of the SG Mix

In [1] the SG Mix was proven to be probabilistically secure under the assumption that the messages arrive according to the Poisson distribution.

In this paper we analyse the two continuous-time mixes under different conditions: we assume MIN/MAX senders and want to see what a locally back-tracing passive adversary might achieve.

Lookback Percentage For a locally back-tracing adversary an important factor is how back in time he will look while enumerating the possible sent messages for a particular delivered message. In the case of the PROB Channel this is obvious (the lookback time should be $\delta_{\max}^{\text{PROB}}$, since earlier messages must have already been delivered), but for the SG Mix this is subject to a decision: it depends on how much of the total infinite past should be considered. With a defined percentage D, the lookback time can be easily calculated from the delay distribution. This lookback time can be thought of as $\delta_{\max}$ for the SG Mix from the adversary's point of view:

$$\begin{aligned} D &= \int_0^{\delta_{\max}^{\text{SG}}} \mu \cdot e^{-\mu\cdot\delta} d\delta = \left[-e^{-\mu\cdot\delta}\right]_0^{\delta_{\max}^{\text{SG}}} = 1 - e^{-\mu\cdot\delta_{\max}^{\text{SG}}} \\ &\Rightarrow \\ \delta_{\max}^{\text{SG}} &= -\frac{1}{\mu} \cdot ln(1-D) \end{aligned} \tag{8}$$

It is clear from the above equation (8) that by increasing the lookback percentage D the lookback time $\delta_{\max}^{\text{SG}}$ approaches infinity logaritmically, but after

a certain limit there is not much point in increasing D, as the efficiency of the algorithm will not increase due to the larger and larger number of messages to be taken into account.

Source-hiding Property for the SG Mix After determining the right look-back percentage (and thus $\delta_{\max}^{\mathrm{SG}}$), local back-tracing could be applied in order to compromise the SG Mix. Our aim is now to apply the general result for the source-hiding property from (7) in order to get an analytical result.

First let us define two abbreviations: $e_{\min} = e^{-\mu \cdot \tau_{\min}}$ and $e_{\max} = e^{-\mu \cdot \tau_{\max}}$. Furthermore, the exponential density function $\mu \cdot e^{-\mu \cdot \delta}$ decreases strictly monotonically, its maximum within an interval is at the start and its minimum at the end. Finally, by applying the value ($\sum_{n=0}^{N-1} r^N = \frac{1-r^N}{1-r}$) for the exponential sums, we can arrive to the following (9):

$$\begin{aligned} \Theta^{\mathrm{SG}} &= \frac{\sum_{i=0}^{\Delta_{\min}-1} e_{\min}^i}{|S| \cdot \sum_{i=0}^{\Delta_{\max}-1} (e_{\max}^i) - 1 + e_{\max}^{\Delta_{\max}}} \\ &= \frac{1}{|S|} \cdot \frac{\frac{1-e_{\min}^{\Delta_{\min}}}{1-e_{\min}}}{\frac{1-e_{\max}^{\Delta_{\max}}}{1-e_{\max}} - 1 + e_{\max}^{\Delta_{\max}}} = \frac{1}{|S| \cdot e_{\max}} \cdot \frac{1-e_{\max}}{1-e_{\min}} \cdot \frac{1-e_{\min}^{\Delta_{\min}}}{1-e_{\max}^{\Delta_{\max}}} \end{aligned} \tag{9}$$

Let us now simplify our equations by introducing the following:

- For the sake of simpicity let us assume that $\delta_{\min} = 0$. This does not significantly change our scenario.
- Let us further assume that $\tau_{\max} = \delta_{\max}$. This is reasonable as one should choose the largest $\tau_{\max}$ in order to make the burden of obligatory message sending for the users as small as possible. (Note that $\tau_{\max}$ cannot be greater than $\delta_{\max}$ in order to be able to *gurantee* the required level of anonymity.)
- Finally, let us denote the ratio $\frac{\tau_{\max}}{\tau_{\min}}$ with K, thus $\tau_{\min} = \frac{\tau_{\max}}{K}$.

With these simplifications we can further evolve the analytical results for the source-hiding property of the SG Mix by:

- $\Delta_{\max} = \lfloor \frac{\delta_{\max} - \delta_{\min}}{\tau_{\max}} \rfloor = \lfloor \frac{\delta_{\max} - 0}{\delta_{\max}} \rfloor = 1$
- $\Delta_{\min} = \lceil \frac{\delta_{\max} - \delta_{\min}}{\tau_{\min}} \rceil = \lceil \frac{\delta_{\max} - 0}{\frac{\tau_{\max}}{K}} \rceil = \lceil \frac{\delta_{\max}}{\frac{\delta_{\max}}{K}} \rceil = \lceil K \rceil \approx K$
- $e_{\max} = e^{-\mu \cdot \tau_{\max}} = e^{-\mu \cdot \delta_{\max}} = e^{-\mu \cdot -\frac{1}{\mu} \cdot ln(1-D)} = 1 - D$
- $e_{\min} = e^{-\mu \cdot \tau_{\min}} = e^{-\mu \cdot \frac{\tau_{\max}}{K}} = e^{-\mu \cdot \frac{\delta_{\max}}{K}} = e^{-\mu \cdot \frac{-\frac{1}{\mu} \cdot ln(1-D)}{K}} = \sqrt[K]{1-D}$

Finally, substituting the above results into (9)we get:

$$\begin{aligned} \Theta^{\mathrm{SG}} &= \frac{1}{|S| \cdot e_{\max}} \cdot \frac{1-e_{\max}}{1-e_{\min}} \cdot \frac{1-e_{\min}^{\Delta_{\min}}}{1-e_{\max}^{\Delta_{\max}}} \\ &\approx \frac{1}{|S| \cdot (1-D)} \cdot \frac{1-(1-D)}{1-\sqrt[K]{1-D}} \cdot \frac{1-(\sqrt[K]{1-D})^K}{1-(1-D)^1} \\ &= \frac{1}{|S| \cdot (1-D)} \cdot \frac{D}{1-\sqrt[K]{1-D}} \end{aligned} \tag{10}$$

With (10) the source-hiding property of the SG Mix is provided. The next step is to look at the PROB Channel and then compare the two analytically.

4.3 Anonymity of the PROB Channel

After having calculated the source-hiding property for the SG Mix, our aim is to do the same for the PROB Channel. Here we have an easier job, since the delay is uniformly distributed. By applying this and some simplifications, (11) gives an approximatation for the source-hiding property of the PROB Channel. The limitation for the equation below is in this case as well $\tau_{\max} \leq \delta_{\max}$.

$$\begin{aligned} \Theta^{\mathrm{PROB}} &= \frac{\sum_{i=1}^{\Delta_{\min}} 1}{|S| \cdot \sum_{i=1}^{\Delta_{\max}} 1} = \frac{\Delta_{\min}}{|S| \cdot \Delta_{\max}} = \frac{\lceil K \rceil}{|S| \cdot 1} = \frac{\lceil K \rceil}{|S|} \\ &\approx \frac{K}{|S|} \end{aligned} \tag{11}$$

4.4 Comparison of the SG Mix and the PROB Channel

After having derived the source-hiding properties of both the SG Mix and the PROB Channel via analytical means, the question now is their ratio Λ. This can be expressed in the following equation:

$$\begin{aligned} \Lambda = \frac{\Theta^{\mathrm{SG}}}{\Theta^{\mathrm{PROB}}} &\approx \frac{\frac{1}{|S| \cdot (1-D)} \cdot \frac{D}{1 - \sqrt[K]{1-D}}}{\frac{K}{|S|}} \\ &\approx \frac{D}{K \cdot (1-D) \cdot (1 - \sqrt[K]{1-D})} \end{aligned} \tag{12}$$

For Λ the following observations can be made:

- If $D > \frac{1}{2}$ then $\Lambda > 1$ regardless of K[4]. This means that with reasonable lookback percentages[5] we will always get a larger source-hiding property for the SG Mix than for the PROB Channel, i.e. the PROB Channel will provide a higher *guaranteed anonymity*.
- For $K \approx 1 \Rightarrow \tau_{\min} \approx \tau_{\max} \Rightarrow \Lambda \approx \frac{1}{1-D} > 1$, which means that if all senders send periodically, then the ratio of the provided anonymity levels will only depend on the lookback percentage.

[4] From its definition naturally $K \geq 1$.

[5] The adversary will choose the largest feasible D, thus normally $D \approx 1$. Our simulations showed that $D \approx 0.95$ yielded the best cost/benefit ratio. Below 0.95 the adversary's success went down, above 0.95 too many messages had to be handled (thus the algorithm got slow) and no real improvement in the back-tracing could be seen.

– Finally, if $D \approx 1$, then $\Lambda >> 1$ meaning that if the adversary chooses a high lookback percentage, i.e. will perform a thorough attack, then the PROB Channel will outperform simply because under such circumstances the SG Mix cannot provide a guaranteed level of anonymity (i.e. $\Theta^{\mathrm{SG}} > 1$)[6].

This result is in contradiction with those presented by Danezis [3]. Two factors may be responsible for this: our traffic model is different (i.e. we used MIN/MAX senders instead of Poisson-distributed messages for the reasons outlined in Section 4.1) and we used a different anonymity metric (the source-hiding property instead on entropy). Our choice of the source-hiding property for measuring anonymity comes from the motivation that we wanted to achieve a system with guaranteed quality of service, i.e. where the system ensures that the required level of anonymity is always achieved. Furthermore, in [9] we already outlined some problems with entropy as a metric for anonymity, which should be further analysed in the light of these new results.

After having presented the analytical results, in the next Section we will introduce some simulation results, which will show the strengths of the two continuous-time systems in practice.

5 Simulation-based Comparison

Unfortunately, the guaranteed levels of anonymity from the previous Section do not give practically usable numbers in several cases. In order to analyse the systems simulations were run to show the performance of both systems in a real scenario.

The simulations were run with MIN/MAX senders and different $\tau_{\min}/\tau_{\max}$ pairs were evaluated. These simulations had the purpose of comparing the two systems: they measured to what extent a locally back-tracing passive adversary might compromise delivered messages. In order to be fair, in all scenarios the mean delay of both the PROB Channel and the SG Mix were the same.

The basic idea of the simulations was that senders send their messages to recipients via the respective AMTS obeying the MIN/MAX rules. Parallel to this the locally back-tracing passive adversary eavesdrops and perceivs the actual message transmissions (to/from the AMTS). Based on these data he then assigns probablities to messages and senders according to (4) and choses the most probable link as his candidate. After the adversary has made his decision, the system checks whether he guessed correctly or not and updates the ratio of correctly matched messages Γ.

It should be noted that the PROB Channel outperformed the SG Mix in almost all the simulated scenarios. This is certainly an interesting result, since

[6] This naturally does not mean that the SG Mix does not provide anonymity at all under such circumstances – it only means that in these cases, based on the source-hiding property, no guaranteed level of anonymity can be ensured, i.e. under extreme conditions messages *could* be traced back to their senders with a probability of 100%.

in [3] Danezis showed that the exponential delay density is optimal (regarding Poisson-distributed incoming messages and an entropy-based anonymity metric).

The following figures show the difference between the SG Mix and the PROB Channel. Fig. 1 shows typical Γ values obtained during the simulation – it can clearly be seen that in the case of the PROB Channel the attacker had a lower success ratio. Fig. 2 shows on the other hand the fraction of the two channels' success ratio, i.e. $\frac{\Gamma^{\mathrm{SG}}}{\Gamma^{\mathrm{PROB}}}$ – again, it is clear that a number above 1.0 depicts the advantage of the PROB Channel[7].

Fig. 1. Γ values for the SG Mix and the PROB Channel

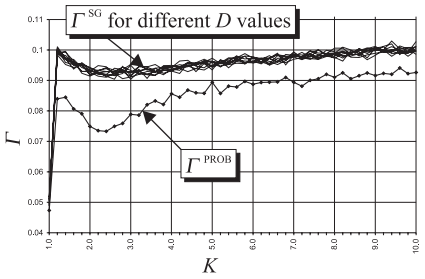
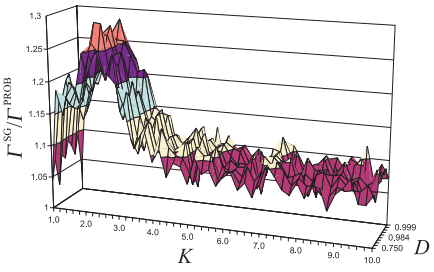


Fig. 2. $\frac{\Gamma^{\mathrm{SG}}}{\Gamma^{\mathrm{PROB}}}$ ratio



For the evaluated scenarios the PROB Channel was on average 12% better than the SG Mix (i.e. the adversary guessed with 12% less certainty the correspondances correctly), while in some cases this advantage went up to 29%. In none of the evaluated cases did the SG Mix beat the PROB Channel.

Although the simulation results turned out less grave as the analytical ones, the PROB Channel outperformed the SG Mix in this case as well. The relatively large difference between the analytical and practical results is mostly due to the fact that with the theoretical source-hiding property we wanted to give an absolute guarantee, i.e. under no circumstances will the anonymity level fall below the given limit, while the simulation results show a global average only.

6 Conclusion

In this paper two continuous-time anonymous message transmission systems – the SG Mix and the PROB Channel – were compared. We analytically derived the maximal guaranteed level of anonymity measured by the source-hiding property for both systems considering a locally back-tracing passive adversary and MIN/MAX senders. We also ran simulations, where we took the ratio of correcly matched messages as the metric for anonymity. In both cases we came to

[7] For both figures $|S| = |R| = 20$, $\delta_{\min} = 0$, $\delta_{\max}^{\mathrm{PROB}} = \tau_{\max} = 20$ and $\delta_{\max}^{\mathrm{SG}}$ was calculated according to (8).

the conclusion that the PROB Channel outperformed the SG Mix contrary to previous work [3]. This result is most probably due to the difference between the methods used for measuring anonymity, a question already touched upon in [9].

Further work is required in order to analyze the different approaches in anonymity metrics and to evaluate both systems under different attacker scenarios. On the other hand, it should be also analysed how these anonymity channels perform if organized into a network and used to transport message streams (e.g. TCP instead of IP).

References

1. Kesdogan, D., Egner, J., Büschkes, R.: Stop-and-go MIXes: Providing probabilistic anonymity in an open system. In: Proceedings of Information Hiding Workshop (IH 1998). Volume 1525 of Springer-Verlag, LNCS., Berkeley, CA (1998)
2. Tóth, G., Hornák, Z.: Measuring anonymity in a non-adaptive, real-time system. In: Proceedings of Privacy Enhancing Technologies (PET2004). Springer-Verlag, LNCS (2004)
3. Danezis, G.: The traffic analysis of continuous-time mixes. In: Proceedings of Privacy Enhancing Technologies (PET2004). Springer-Verlag, LNCS (2004)
4. Danezis, G., Laurie, B.: Minx: A simple and efficient anonymous packet format. In: Proceedings of the Workshop on Privacy in the Electronic Society, Washington DC (2004)
5. Berthold, O., Federrath, H., Köpsell, S.: Web mixes: A system for anonymous and unobservable internet access. In Federrath, H., ed.: Designing Privacy Enhancing Technologies. Volume 2009 of Springer-Verlag, LNCS., Berkeley, CA (2001) 115–129
6. Pfitzmann, A., Köhntopp, M.: Anonymity, unobservability, and pseudonymity – a proposal for terminology. In: Designing Privacy Enhancing Technologies. Volume 2009 of Springer-Verlag, LNCS., Berkeley, CA (2001) 1–9
7. Serjantov, A., Danezis, G.: Towards an information theoretic metric for anonymity. In: Proceedings of Privacy Enhancing Technologies (PET2002). Volume 2482 of Springer-Verlag, LNCS., San Francisco, CA (2002)
8. Díaz, C., Seys, S., Claessens, J., Preneel, B.: Towards measuring anonymity. In: Proceedings of Privacy Enhancing Technologies (PET2002). Volume 2482 of Springer-Verlag, LNCS., San Francisco, CA (2002) 54–68
9. Tóth, G., Hornák, Z., Vajda, F.: Measuring anonymity revisited. In Liimatainen, S., Virtanen, T., eds.: Proceedings of the Ninth Nordic Workshop on Secure IT Systems, Espoo, Finland (2004) 85–90
10. Chaum, D.: Untraceable electronic mail, return addresses, and digital pseudonyms. Communications of the ACM **4** (1981) 84–88
11. Díaz, C., Sassaman, L., Dewitte, E.: Comparison between two practical mix designs. In: Proceedings of ESORICS: 9th European Symposium on Research in Computer Security. Volume 3193 of Springer-Verlag, LNCS., French Riviera, France (2004) 141–159
12. Tóth, G., Hornák, Z.: The aprob-channel: adaptive semi-real-time anonymous communication. In: Security and Privacy in Dynamic Environments – Proceedings of the IFIP TC-11 21st International Information Security Conference (SEC 2006). Volume 201 of Springer-Verlag, IFIP International Federation for Information Processing., Karlstad, Sweden (2006)