

Anonymous communication

Claudia Díaz and Bart Preneel

K.U.Leuven Dept. Electrical Engineering-ESAT/COSIC
Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium
`claudia.diaz@esat.kuleuven.ac.be`, `bart.preneel@esat.kuleuven.ac.be`
<http://www.esat.kuleuven.ac.be/cosic/>

Abstract. This paper presents an overview on anonymous communication systems. It describes the different types of mixes and how they can be combined in mix networks in order to implement an anonymous communication infrastructure. Also, it gives some notions on how to measure the anonymity provided to the users by an anonymous system. Dummy traffic is introduced as a tool to enhance the anonymity of the system and the robustness against attacks. The convenience of introducing controls in an anonymous communication system is discussed. This paper also indicates the main open research areas related to anonymous communication systems.

1 Introduction

The Internet was initially perceived as a rather anonymous environment. Nowadays, we know that it is a powerful surveillance tool: anyone willing to listen to the communication links can spy on you, and search engines and data mining techniques are becoming increasingly powerful. Privacy does not only mean confidentiality of the information; it also means not revealing information about who is communicating with whom.

When a user accesses the Internet through a non-anonymous system, he is leaking information such as the user's computer IP address. This means that all his actions become linkable (which allows detailed profiling of the user) and this information could lead to an identification of the user in the off-line world. Note that this information can be gathered without the consent, or even knowledge of the user. Therefore, anonymity needs to be implemented at the communication and application layer in order to effectively protect the users' privacy. A user that connects to the Internet through anonymous communication systems protects his privacy towards the other end of the communication line and towards third parties.

2 Building blocks for anonymous communication

Mixes are the most popular basic building block used to implement anonymous communications. Other anonymous communication systems have been implemented using peer-to-peer models [FM02,RR98,GRPS03,SBS02,BG03] or multicast protocols [SL00].

In this section we focus on mix networks. We first introduce the concept of a mix, and then describe the different types of mixes that we encounter in the literature. Finally, we discuss the different topologies with which mixes may be combined to configure mix networks.

2.1 What is a mix?

Mixes were proposed by Chaum [Cha81] in 1981. The mix takes a number of input messages, and outputs them in such a way that it is not possible to link an output to the corresponding input. In order to achieve this goal, the mix changes the appearance (by encrypting and padding messages) and the flow of messages (by delaying and reordering). The technique used by Chaum's mix to change the flow of messages was to collect n inputs, shuffle and output them.

The functionality of this first design was to provide untraceable email, but mixes may be used for providing anonymity to a wide range of applications (e.g., the Onion Routing system [GRS96] offers application-independent connection anonymity; a working anonymous web browsing system based on mixes is JAP [JAP]). Instead of mixing email messages, we can make abstraction of the type of information contained in the messages. Note that some mix designs are not appropriate for applications that have strong real-time requirements, because the delay introduced by the mixing may make the system impractical.

Other work on mixes focussed on the correctness and verifiability of the decriptions and the mixing algorithm has been developed, among others, by Jakobsson [Jak98] and Abe [Abe98].

Pool mixes Many variations to Chaum's original design have been proposed in the last years. First, a *pool* was added to the mix. In this design, the mix does not output all the messages it contains; instead, it keeps a certain number of them to be mixed with new inputs. Also, timeouts have been proposed as the mechanism to trigger the flushing of messages. Several strategies can be used in order to optimize the anonymity service provided by the mix by playing with the different parameters. Díaz and Serjantov propose in [DS03b] a model that generalizes pool mixes.

Continuous mixes A different mix concept was proposed by Kesdogan *et al.* in [KEB98]. In this design, the messages are delayed a certain amount of time, chosen by the user from an exponential distribution. The advantage of this system is that the delay does not depend on the traffic that arrives to the mix. On the other hand, the anonymity provided to the users may go to low levels if the number of users decreases during a certain period of time.

2.2 Mix network topologies

In order to increase the anonymity of a mix system, mixes are usually combined in a mix network. This way, the fact of some mixes being corrupted or controlled

by an attacker does not break the anonymity of the users (the anonymity of a message is guaranteed even if only one of the mixes in the path of the message is honest). Also, the reliability of the system is improved, because the failure of a mix does not lead to a denial of service.

The two classical types of mix network that have been considered are cascades and free route networks. In a cascade, the possible paths that a message can follow are predefined; in a free route network, users select freely their own path, which may be different for every message. The advantages and disadvantages of these two topologies have been pointed out by Berthold *et al.* in [BPS00].

More recently, Danezis proposed in [Dan03] a mix network topology that is somehow in between the two classical designs. In this model, every mix node communicates with a few neighboring others. The goal of this idea is to combine the advantages of cascades and free route networks and overcome the disadvantages.

3 Anonymity metrics

How to measure the degree of anonymity offered to the users of a mix network? An attacker may deploy passive attacks (i.e., traffic analysis [SS03]) or active attacks (e.g., the *blending* or *n-1* attack, described by Serjantov *et al.* in [SDS02]) in order to identify the sender (or recipient) of a message.

The attacker may normally obtain probabilistic relationships between the inputs and the outputs of a mix. In certain conditions (for example, low traffic, or active attacks), the attacker may be able to narrow down the set of possible senders (recipients) of a message. In other cases, one of the users will appear as having a very high probability of being the sender of a particular message.

Based on the definition for anonymity proposed by Pfitzmann and Köhntopp in [PK00], two information theoretic models were proposed by Díaz *et al.* in [DSCP02] and by Serjantov and Danezis in [SD02]. These models measure the anonymity provided by a mix towards an attacker, whose powers must be clearly specified before applying the anonymity metric. The anonymity is measured using the concept of *entropy* (i.e., uncertainty), taking into account the probabilistic information that an attacker is able to obtain from the system.

The metrics may be applied to measure the uncertainty of the attacker about the sender of the message, i.e., *sender anonymity*. Analogously, the uncertainty of the attacker regarding the recipient of a message, i.e., *recipient anonymity* may be computed.

4 Dummy traffic

A dummy message is a *fake* message introduced in the network in order to make it more difficult for an attacker to deploy traffic analysis attacks. The dummy message is normally produced by the mixes, and they have as destination another mix, instead of a real recipient. Dai proposed in Popenet [Dai96] a system in which

the traffic is constant: the links between mixes are padded with dummy messages whenever the real traffic is not enough to fill them. This system provides not only anonymity, but also unobservability, since an observer of the network cannot tell whether there are real messages traveling in the network or not. Unfortunately, the system is not practical.

The generation and transmission of dummy traffic has a cost, and it is therefore very important to find the right balance on the amount of dummies that should be created in a mix network. Also, some other questions regarding dummy traffic remain open, such as if it should or should not depend on the real traffic load or what is the most appropriate route length for the dummy messages. Some of these aspects are discussed by Díaz and Preneel in [DP03].

Dummy traffic may also be used to detect active attacks on a mix network, as described by Danezis and Sassaman in [DS03a].

5 Controlled anonymous communication?

One of the obstacles for the wide deployment of anonymous communication infrastructures is the fear of governments, businesses and institutions of not being able to control abuse and crime. Nevertheless, it is worth noting that malicious users do often use other means than anonymity to abuse systems, as identity theft, for example.

The introduction of control techniques in anonymous systems also helps compliance with legal requirements, such as accountability. Note also that some applications, such as electronic commerce would not be trustful for users and vendors if control mechanisms to prevent fraud are not implemented. A model for anonymity control mechanisms is presented in [CDN⁺03].

The deployment of a controlled anonymous infrastructure seems to meet government and users requirements. In such a system, honest users remain anonymous, but misbehaving users may be identified with the collaboration of a trusted third party. Claessens *et al.* have proposed a controlled anonymous communication system in [CDG⁺03] based on cryptographic techniques such as fair blind signatures [AO01,SPC95] or group signatures [ACJT00].

6 Conclusions and future work

This paper presents the state-of-the-art regarding anonymous communication systems. We have first introduced the concept of a mix and described the main different variants to the original design. The necessity of combining several mixes in a network has been justified, and the different possible topologies of the network have been discussed.

We have given an introduction to the techniques that can be used to measure the degree of anonymity provided by anonymous communication systems. Dummy traffic has been introduced as means to increase the anonymity and the robustness of the system against certain attacks.

Finally, we have discussed the need of introducing controls in anonymous systems in order to prevent misuse.

There are still many issues that need further work in order to have a good understanding of anonymous communication systems. We point out the following:

- The current anonymity metrics can measure the anonymity provided by a mix in a simulation or in a working setting, but we do not have yet theoretical tools that allow us to know the anonymity properties of the mix during the design phase.
- The anonymity metrics are very useful to measure the anonymity provided by a single mix, but they fail to measure the end-to-end anonymity provided by a mix network. An extension to the metric needs to be found in order to have practical tools to measure the anonymity provided by a mix network.
- In the presence of dummy traffic, the sender anonymity cannot be computed with the intuitive extension of the metric. A generalization of the metric for the use of dummy traffic should solve this problem.
- Much research need to be done in order to solve many dummy traffic related problems. We do not know yet which is the most appropriate distribution for the generation of dummies, the route length they should have in order to optimize the cost/anonymity relationship, whether they should be inserted in the pool of the mix or at the output, whether dummy traffic should depend on the real traffic traveling in the network or not, and how this dependency should be.
- Different mix designs need to be compared in order to find the best performing mixes.
- Regarding anonymity control, not much work has been done in the area of anonymous communications. We still lack practical designs that implement anonymity control and that are efficient at the same time.

Acknowledgments

Claudia Díaz is funded by a research grant of the K.U.Leuven. This work was also partially supported by the IWT STWW project on Anonymity and Privacy in Electronic Services (APES), and by the Concerted Research Action (GOA) Mefisto-2000/06 of the Flemish Government.

References

- [Abe98] Masayuki Abe. Universally verifiable mix-net with verification work independent of the number of mix-servers. In *Proceedings of EUROCRYPT'98*, pages 437–447. Springer-Verlag, LNCS 1403, 1998.
- [ACJT00] Giuseppe Ateniese, Jan Camenisch, Marc Joye, and Gene Tsudik. A practical and provably secure coalition-resistant group signature scheme. *Lecture Notes in Computer Science*, 1880:255+, 2000.

- [AO01] Masayuki Abe and Miyako Ohkubo. Provably Secure Fair Blind Signatures with Tight Revocation. In Colin Boyd, editor, *Advances in Cryptology – ASIACRYPT 2001*, Lecture Notes in Computer Science, LNCS 2248, pages 583–601. Springer-Verlag, December 2001.
- [BG03] Krista Bennett and Christian Grothoff. GAP – practical anonymous networking. In Roger Dingledine, editor, *Proceedings of Privacy Enhancing Technologies workshop (PET 2003)*. Springer-Verlag, LNCS 2760, March 2003.
- [BPS00] Oliver Berthold, Andreas Pfitzmann, and Ronny Standtke. The disadvantages of free MIX routes and how to overcome them. In H. Federrath, editor, *Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*, pages 30–45. Springer-Verlag, LNCS 2009, July 2000.
- [CDG⁺03] Joris Claessens, Claudia Diaz, Caroline Goemans, Bart Preneel, Joos Vandewalle, and Jos Dumortier. Revocable anonymous access to the internet. *Journal of Internet Research: Electronic Networking Applications and Policy*, 13(4):242–258, 2003.
- [CDN⁺03] Joris Claessens, Claudia Diaz, Svetla Nikova, Vincent Naessens, Bart de Win, Caroline Goemans, Stefaan Seys, Mieke Loncke, Jos Dumortier, Bart de Decker, and Bart Preneel. Apes deliverable d7: Applications requirements for controlled anonymity. Technical report, K.U.Leuven, May 2003.
- [Cha81] David Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 4(2), February 1981.
- [Dai96] Wei Dai. PIPenet 1.1. Usenet post, August 1996.
- [Dan03] George Danezis. Mix-networks with restricted routes. In Roger Dingledine, editor, *Proceedings of Privacy Enhancing Technologies workshop (PET 2003)*. Springer-Verlag, LNCS 2760, March 2003.
- [DP03] Claudia Diaz and Bart Preneel. Reasoning about the anonymity provided by pool mixes that generate dummy traffic. Technical report, K.U.Leuven, December 2003.
- [DS03a] George Danezis and Len Sassaman. Heartbeat traffic to counter (n-1) attacks. In *Proceedings of the Workshop on Privacy in the Electronic Society (WPES 2003)*, Washington, DC, USA, November 2003.
- [DS03b] Claudia Diaz and Andrei Serjantov. Generalising mixes. In Roger Dingledine, editor, *Proceedings of Privacy Enhancing Technologies workshop (PET 2003)*. Springer-Verlag, LNCS 2760, March 2003.
- [DSCP02] Claudia Diaz, Stefaan Seys, Joris Claessens, and Bart Preneel. Towards measuring anonymity. In Roger Dingledine and Paul Syverson, editors, *Proceedings of Privacy Enhancing Technologies Workshop (PET 2002)*. Springer-Verlag, LNCS 2482, April 2002.
- [FM02] Michael J. Freedman and Robert Morris. Tarzan: A peer-to-peer anonymizing network layer. In *Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS 2002)*, Washington, DC, November 2002.
- [GRPS03] Sharad Goel, Mark Robson, Milo Polte, and Emin Gun Sirer. Herbivore: A Scalable and Efficient Protocol for Anonymous Communication. Technical Report 2003-1890, Cornell University, Ithaca, NY, February 2003.
- [GRS96] David M. Goldschlag, Michael G. Reed, and Paul F. Syverson. Hiding Routing Information. In R. Anderson, editor, *Proceedings of Information Hiding:*

- First International Workshop*, pages 137–150. Springer-Verlag, LNCS 1174, May 1996.
- [Jak98] Markus Jakobsson. A practical mix. In *Proceedings of EUROCRYPT'98*, pages 448–461. Springer-Verlag, LNCS 1403, 1998.
- [JAP] JAP Anonymity & Privacy. <http://anon.inf.tu-dresden.de/>.
- [KEB98] Dogan Kesdogan, Jan Egner, and Roland Büschkes. Stop-and-go MIXes: Providing probabilistic anonymity in an open system. In *Proceedings of Information Hiding Workshop (IH 1998)*. Springer-Verlag, LNCS 1525, 1998.
- [PK00] Andreas Pfitzmann and Marit Kohntopp. Anonymity, unobservability and pseudonymity — a proposal for terminology. In H. Federrath, editor, *Designing Privacy Enhancing Technologies: Proceedings of the International Workshop on the Design Issues in Anonymity and Observability*, pages 1–9. Springer-Verlag, LNCS 2009, July 2000.
- [RR98] Michael Reiter and Aviel Rubin. Crowds: Anonymity for web transactions. *ACM Transactions on Information and System Security*, 1(1), June 1998.
- [SBS02] Rob Sherwood, Bobby Bhattacharjee, and Aravind Srinivasan. P5: A protocol for scalable anonymous communication. In *Proceedings of the 2002 IEEE Symposium on Security and Privacy*, May 2002.
- [SD02] Andrei Serjantov and George Danezis. Towards an information theoretic metric for anonymity. In Roger Dingledine and Paul Syverson, editors, *Proceedings of Privacy Enhancing Technologies Workshop (PET 2002)*. Springer-Verlag, LNCS 2482, April 2002.
- [SDS02] Andrei Serjantov, Roger Dingledine, and Paul Syverson. From a trickle to a flood: Active attacks on several mix types. In Fabien Petitcolas, editor, *Proceedings of Information Hiding Workshop (IH 2002)*. Springer-Verlag, LNCS 2578, October 2002.
- [SL00] Clay Shields and Brian Neil Levine. A protocol for anonymous communication over the internet. In *ACM Conference on Computer and Communications Security*, pages 33–42, 2000.
- [SPC95] Markus A. Stadler, Jean-Marc Piveteau, and Jan L. Camenisch. Fair blind signatures. *Lecture Notes in Computer Science*, 921:209+, 1995.
- [SS03] Andrei Serjantov and Peter Sewell. Passive attack analysis for connection-based anonymity systems. In *Proceedings of ESORICS 2003*, October 2003.