Anonymity Metrics Revisited

Claudia Díaz, *COSIC (K.U.Leuven)*
claudia.diaz@esat.kuleuven.be

***Abstract*— In 2001, two information theoretic anonymity metrics were proposed: the *effective anonymity set size* and the *degree of anonymity*. Here, we propose an abstract model for a general anonymity system which is consistent with the definition of anonymity on which the metrics are based. We revisit entropy-based anonymity metrics, and we apply them to Crowds, a practical anonymity system. We discuss the differences between the two metrics and the results obtained in the example.**

***Index Terms*— Anonymity, metrics, entropy**

I. Introduction

The need of a metric to measure the performance of anonymity implementations appeared with the development of applications that enabled anonymous electronic transactions, such as untraceable email, electronic voting, anonymous e-coins or privacy-enhanced web browsing.

The research questions that arose were: how can anonymity be measured? How can two different anonymity systems be compared? Is there a general anonymity metric which can be applied to any anonymity system? How can we evaluate the effectiveness of different attacks on the anonymity system? How can we quantify losses and gains in anonymity? How can anonymity metrics reflect the partial or statistic information often obtained by an adversary? The metrics described here provide answers to these questions.

The anonymity metrics presented in this paper were originally proposed in [3], [9], and can be applied to concrete systems, adversaries, and conditions. These metrics give a measure of the size and distinguishability of the set of subjects potentially linked to a particular transaction, and attacked by a concrete adversary. In order to get an idea on the performance of an anonymity implementation under different conditions, multiple anonymity measurements must be made and analyzed.

Information theoretic metrics can be applied to a broad range of anonymity systems. It is thus important to understand the concepts behind entropy-based anonymity metrics in order to apply and interpret them correctly in concrete scenarios. The metrics must be adapted to the anonymity system under study, and the computation of probability distributions that lead to meaningful metric values is not always obvious.

We put this work into context by describing the related work in Sec. II. The model for anonymity systems is described in Sect III, and the attack model in Sect. IV. Section V describes information theoretic anonymity metrics, which are then applied to a practical example in Sect VI. Finally, Sect. VII presents the conclusions of this paper.

II. Related work

A. Defining anonymity

Prior to the quantification of anonymity, a working definition for the term *anonymity* was needed. Pfitzmann and Hansen [7] defined *anonymity* as *the state of being not identifiable within a set of subjects, the anonymity set*. This definition, first proposed in year 2000, has been adopted in most of the anonymity literature.

According to the Pfitzmann-Hansen definition of anonymity, the subjects who may be related to an anonymous transaction constitute the *anonymity set* for that particular transaction. A subject carries on the transaction *anonymously* if he cannot be distinguished (by an adversary) from other subjects. This definition of anonymity captures the probabilistic information obtained by adversaries trying to identify anonymous subjects, as we explain in Sect. V.

B. Anonymity metrics

Before information theoretic anonymity metrics were proposed, there had been some attempts to quantify anonymity in communication networks.

Reiter and Rubin [8] define the *degree of anonymity* as a probability $1 - p$, where p is the probability assigned by an attacker to potential senders. In this model, users are more anonymous as they appear (towards a certain adversary) to be less likely of having sent a message. This metric considers users separately, and therefore does not capture anonymity properties very well. Consider a first system with 2 users which appear to be the sender of a message with probability $1/2$. Now consider a second system with 1000 users. User u_1 appears as the sender with probability $1/2$, while all the other users are assigned probabilities of having sent the message below 0.001. According to the definition of Reiter and Rubin, the *degree of anonymity* of u_1 and of the two users of the first system would be the same (50%). However, in the second system, u_1 looks much more likely to be the sender than any other user; while the two users of the first system are indistinguishable to the adversary.

Berthold *et al.* [1] define the *degree of anonymity* as $A = log_2(N)$, where N is the number of users of the system. This metric only depends on the number of users of the system, and therefore does not express the anonymity properties of different systems. Moreover, adversaries may be able to obtain probabilistic information on the set of potential senders, which is not taken into account in this metric.

Information theoretic anonymity metrics were independently proposed in two papers presented at the *2nd Workshop on Privacy Enhancing Technologies*. The basic principle of both metrics is the same. The metric proposed by Serjantov

Dagstuhl Seminar Proceedings 05411
Anonymous Communication and its Applications
http://drops.dagstuhl.de/opus/volltexte/2006/483

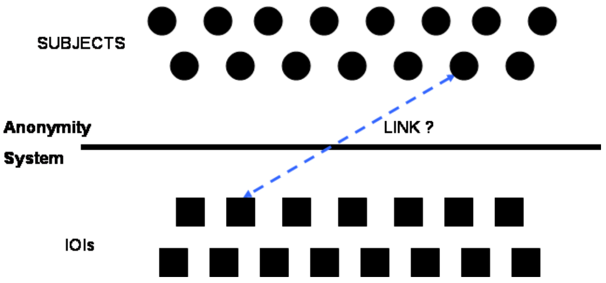

Fig. 1. Model for Anonymity Systems

and Danezis [9] uses entropy as measure of the *effective anonymity set size*. The one presented by Díaz *et al.* [3] goes one step further, normalizing the entropy to obtain a *degree of anonymity* in the scale $0..1$. The details of the two flavors of anonymity metrics are explained in Sect. V.

Examples on how to apply information theoretic anonymity metrics to practical anonymous communication systems based on mixes have been presented by Díaz *et al.* in [4], [5], [6].

III. Model

Many anonymity systems can be modeled in terms of unlinkability. Unlinkability is defined by Pfitzmann and Hansen [7] as follows:*unlinkability of two or more items means that within this system, these items are no more and no less related than they are related concerning the a priori knowledge*.

Our model can be applied both to *sender* and *recipient anonymity*. If we consider the sending and receiving of messages as *Items Of Interest* (IOIs), anonymity may be defined as unlinkability of an IOI and a subject. More specifically, we can describe the anonymity of an IOI such that it is not linkable to any subject, and the anonymity of a subject as not being linkable to any IOI. In this context, *unlinkability* is achieved with high entropy values.

Figure 1 presents a simplified anonymity model. The goal of anonymity systems is to hide the relationship between subjects and IOIs. Hiding these links is the basic mechanism behind anonymous transactions.

An observer of the system sees that a set of users are accessing the anonymity system. At the output of the system, they see IOIs which are hard to link to a particular subject. The set of subjects who might be linked to an IOI is called the *anonymity set*. The larger the *anonymity set*, the more anonymity a subject is enjoying. The notion of *anonymity set* is key to define anonymity metrics, as we show in Sect. V-B.

IV. Attack model

We can distinguish two types of attacks on anonymity systems: *attacks on anonymity* and *attacks on the availability* of the anonymity service (also called *denial of service attacks*). Denial of service attacks may only be deployed by *active* attackers (see description below). These attacks are aimed at reducing the availability of the system, which may be a goal in itself, or part of an attack on anonymity (e.g., the adversary may block several entities from accessing the system in order to reduce the anonymity set). In this paper, we are interested in the effects of the attacks on anonymity. More specifically, in measuring the certainty of the adversary on the existence of a link between a subject and an IOI.

The quantification of anonymity is dependent on the *adversary* or *attacker* considered. The adversary has certain capabilities and deploys attacks in order to gain information and find links between subjects and IOIs. Most of these attacks lead to a distribution of probabilities that assign users a certain probability of being linked (either as senders or as recipients) to IOIs.

The metric we propose here takes into account the probabilities assigned by the adversary to users potentially linked to an IOI. Note that the metric measures anonymity *with respect to* a particular attack; results are no longer valid if the attack model changes. Therefore, concrete assumptions about the attacker have to be clearly specified when measuring anonymity. Some of the adversary's properties we should make explicit are [1]:

- *Passive-Active:* A passive attacker listens to the communication and/or reads internal information of entities participating in the protocols, passive attackers typically perform traffic analysis of the communication. Active attackers can add, remove or modify messages and adapt internal information of participating entities.
- *Internal-External:* An internal attacker controls one or several entities that are part of the system (e.g., the attacker controls communication nodes). External attackers only control communication links.
- *Partial-Global:* A global attacker has access to the entire communication system (e.g., all communication links), while a partial attacker (also called *local attacker* in the literature) only sees part of the resources (e.g., a limited number of peers in a peer-to-peer network).
- *Static-Adaptive:* Static attackers control a predefined set of resources and are unable to alter their behavior once a transaction is in progress. Adaptive attackers gain control on new resources or modify their behavior, depending on intermediate results of the attack.
- *Temporary-Permanent:* Permanent adversary have been observing the system since it started functioning and knows its whole history. Temporary attackers start observing or attacking the system at time t_0, and they do not have information on events previous to t_0.

V. Information theoretic anonymity metrics

In this section, we first introduce the concept of entropy, on which information theoretic anonymity metrics are based. Then, we explain how the *effective anonymity set size* and the *degree of anonymity* can be computed. The metrics presented in this section are applied to a practical system in Sect. VI.

A. Entropy

The information theoretic concept of entropy [10] provides a measure of the uncertainty of a random variable. Let X be the discrete random variable with probability mass function $p_i = Pr(X = i)$, where i represents each possible value that

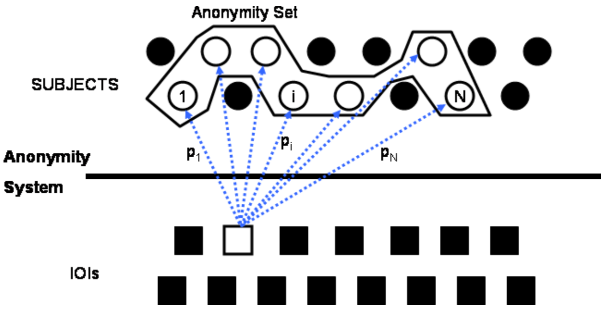

Fig. 2. Anonymity set

X may take with probability $p_i > 0$. In this case, each i corresponds to a subject of the anonymity set; i.e., p_i is the probability of subject i being linked to the IOI.

We denote by $H(X)$ the entropy of a random variable, and by N the number of subjects in the anonymity set. $H(X)$ can be calculated as:

$$H(X) = -\sum_{i=1}^{N} p_i \log_2(p_i) \ .$$

B. Effective anonymity set size

The *effective anonymity set size* is an intermediate step to compute the *degree of anonymity*. Serjantov and Danezis proposed in [9] the use of the effective anonymity set size as metric.

As mentioned in Sect. II-A, *anonymity* was defined by Pfitzmann and Hansen [7] as *the state of being not identifiable within a set of subjects, the anonymity set*. Anonymity metrics aim at giving a meaningful measure of the anonymity set size.

After deploying an attack on an anonymity system, the adversary typically obtains a distribution of probabilities that link subjects to the particular IOI of the attack. The probabilities are shown in Fig. 2 with the arrows that connect the IOI to the subjects of the anonymity set. Different subjects may appear as having a higher or lower probability p_i of link with the IOI, depending of the information obtained by the adversary using the attack.

Let N be the total number of subjects which are linked to the IOI with a non-zero probability ($p_i > 0$, $i = 1..N$). The *effective anonymity set size* is defined as the entropy $H(X)$ of the distribution X of probabilities that link the subjects of the anonymity set to the IOI.

Entropy-based anonymity metrics give a measure of the uncertainty of the adversary on the subject who is related to the IOI. The *effective anonymity set size* takes into account the *number* of potential subjects linked to the IOI, and the *probabilities* assigned to the subjects.

The metric (and thus anonymity) increases its value with two factors. First, with the number of subjects potentially linked to the IOI; and second, with the uniformity of the probability distribution. The more equally distributed the probabilities assigned to the subjects of the anonymity set, the higher the entropy (i.e., the higher the effective anonymity set size).

C. Degree of anonymity

The *degree of anonymity* is a normalized version of the *effective anonymity set size*, which tells tells how good the system is performing on a $0-1$ scale. This metric is an original contribution of Díaz *et al.* and was proposed in [3] (note that both metrics were proposed independently at the same time).

The maximum effective anonymity set size for N subjects is reached when all subjects are linked to the IOI with equal probability (i.e., $p_i = 1/N$). In this case, all subjects are indistinguishable towards the adversary with respect to the IOI. For a given number N of users, the maximum achievable anonymity corresponds to the entropy of a uniform distribution. We denote the maximum entropy by H_M:

$$H_M = \log_2(N) \ .$$

If we assume that the adversary has no *a priori* information on the system (i.e., the *a priori* anonymity of an IOI is H_M), the amount of information gained by the adversary with an attack is the difference in the entropy before and after the attack, that is: $H_M - H(X)$.

The *degree of anonymity* is defined as the normalized value of this difference in knowledge of the adversary:

$$d = 1 - \frac{H_M - H(X)}{H_M} = \frac{H(X)}{H_M} \ .$$

As we can observe in the formula, the *degree of anonymity* is obtained dividing the *effective anonymity set size* by the maximum entropy for a given number of subjects. This *degree* evaluates how much anonymity is provided by a system independently from the number of users. Given a certain number of subjects, the computation of the *degree of anonymity* gives an idea on how close the anonymity of the subjects is to the maximum achievable.

Both metrics are computed using the same information, and one can trivially be computed from the other. The difference is, however, that the *effective anonymity set size* ties the anonymity to the actual number of users in the system; while the *degree of anonymity* makes abstraction on the number of users and focusses on the performance of the system (i.e., how close it is to the maximum achievable anonymity).

VI. Example: Crowds

Crowds [8] is designed to provide anonymity to users who want to access web pages. To achieve this goal, the designers introduce the notion of *blending into a crowd*: users are grouped into a set, and they forward requests within this set before the request is sent to the web server. The web server cannot know from which member the request originated, since it gets the request from a random member of the crowd, who is forwarding the message on behalf of the real originator. The users (members of the crowd) are called *jondos*.

The system works as follows: when a *jondo* wants to request a web page it sends the request to a second (randomly chosen) *jondo*. This *jondo* will, with probability p_f, forward the request to a third *jondo* (again, randomly chosen), and will, with probability $(1 - p_f)$ submit it to the server. Each *jondo* in the path (except for the first one) chooses to forward or submit the

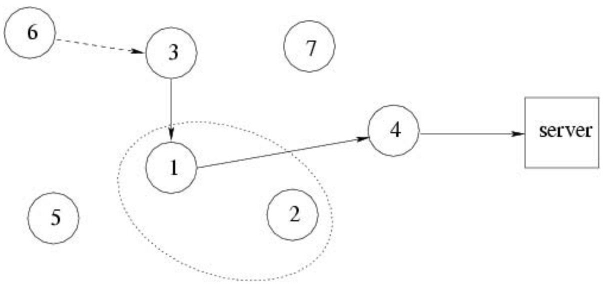

Fig. 3. Example of a Crowds system with 7 *jondos*

request *independently* from the decisions of the predecessors in the path.

Communication between *jondos* is encrypted, and the final request to the server is sent in clear text. Every *jondo* can observe the contents of the message (and thus the address of the target server), but it cannot know whether the predecessor is the originator of the message or whether he is just forwarding a message received by another member.

A. Attack Model

In this section we calculate the degree of anonymity provided by Crowds to its users, with respect to colluding crowd members, that is, a *set of corrupted jondos that collaborate in order to disclose the identity of the jondo that originated the request*. The assumptions made on the attacker are:

- *Internal*: The attacker controls some of the entities which are part of the system.
- *Passive*: The corrupted *jondos* can listen to communication. Although they have the ability to add or delete messages, they do not gain extra information on the identity of the originator by doing so.
- *Partial*: We assume the attacker controls a limited set C of *jondos*, and cannot perform any traffic analysis on the rest of the system.
- *Static*: The set of jondos controlled by the adversary is fixed.
- *Temporary*: The adversary does not need to observe the system for long time to deploy this attack. Permanent attackers may refine the attack by correlating subsequent connections.

B. Effective anonymity set size

Figure 3 shows an example of a crowds system. In this example the *jondos* 1 and 2 are controlled by the attacker, i.e., they are *colluding crowd members*. An honest *jondo* creates a path that includes at least one corrupted *jondo*.[1] The adversary wants to know which of the *jondos* is the real originator of the message.

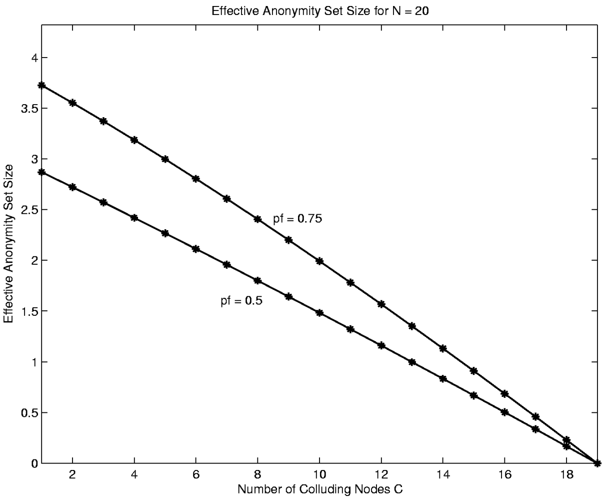

Fig. 4. Effective Anonymity Set Size for Crowds (N=20)

In a general Crowds network, let N denote the number of members of the crowd, C the number of malicious collaborators, p_f the probability of forwarding and p_i the probability of being originator of a request assigned by the attacker to *jondo i*. From [8] we know that, under the described attack model, the probability assigned to the predecessor of the first malicious *jondo* in the path (for simplicity, let this *jondo* be number *C+1*) equals:

$$p_{C+1} = \frac{N - p_f(N-C-1)}{N} = 1 - p_f \frac{N-C-1}{N} \quad .$$

The probabilities assigned to the colluding *jondos* remain zero, and assuming that the adversary does not have any extra information about honest nodes, the probabilities assigned to those members are:

$$p_i = \frac{1 - p_{C+1}}{N-C-1} = \frac{p_f}{N} \quad , \quad C+2 \leq i \leq N \quad .$$

Applying the formula of the entropy presented in Sect. V-A, the effective anonymity set size under these attack conditions can be computed as:

$$H(X) = \frac{N - p_f(N-C-1)}{N} \log_2 \left[\frac{N}{N - p_f(N-C-1)} \right] + p_f \frac{N-C-1}{N} \log_2 \left[\frac{N}{p_f} \right] \quad .$$

As we can see, the *effective anonymity set size* for Crowds is a function of N, C and p_f. In order to show the variation of $H(X)$ with respect to these parameters we chose $p_f = 0.5$ and $p_f = 0.75$. The effective anonymity set sizes for a system with $N = 20$ and $N = 100$ are shown in Fig. 4 and Fig. 5. The anonymity metric is a function of the number C of colluding *jondos*, which takes values between 1 and $N-1$. Note that if $C = 0$ there is no adversary, and the effective anonymity set size is maximum ($\log_2(N)$); if $C = N$ the adversary controls all *jondos*, leaving none to attack.

As we can see in Fig. 4 and Fig. 5, the effective anonymity set size decreases almost linearly with the number of colluding *jondos* (controlled by the adversary), down to zero when the adversary controls $N-1$ *jondos* (and is thus able to uniquely identify messages sent by the remaining *jondo*). We can also see in the figures that the effective anonymity set size

[1] If the path does not go through a corrupted *jondo* the attacker cannot get any information.

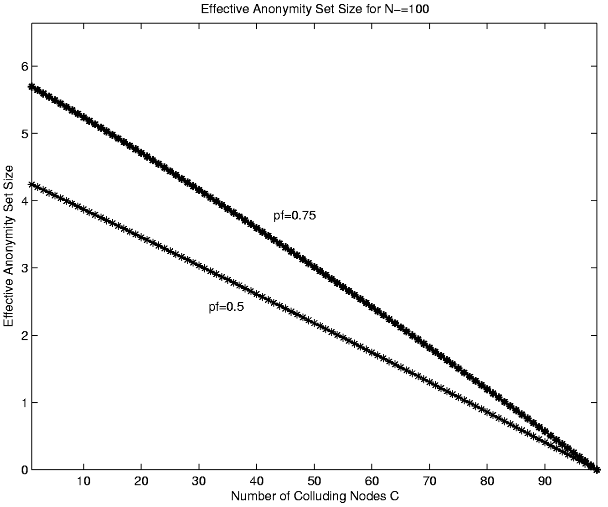

Fig. 5. Effective Anonymity Set Size for Crowds (N=100)

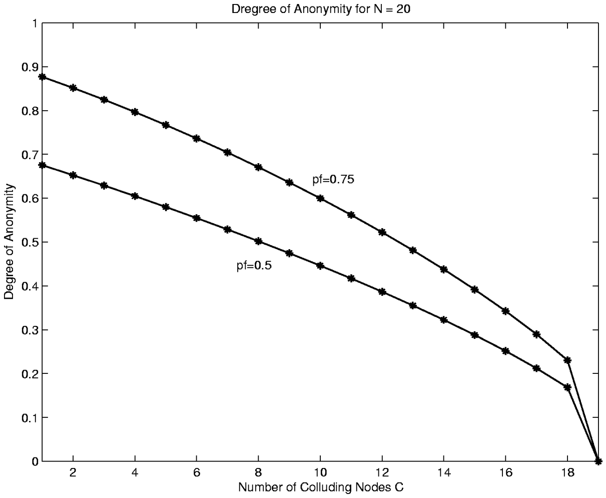

Fig. 6. Degree of Anonymity for Crowds (N=20)

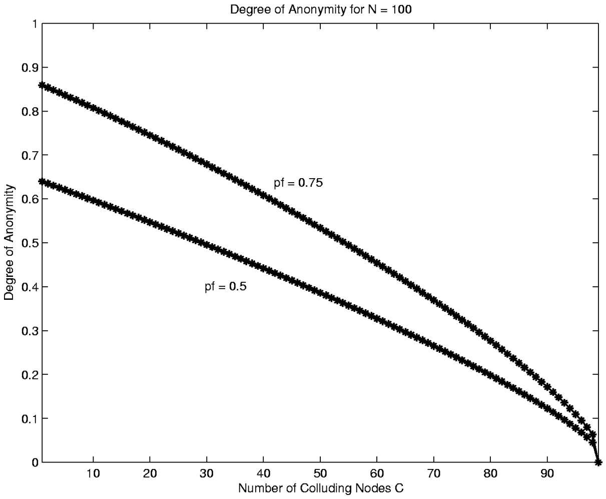

Fig. 7. Degree of Anonymity for Crowds (N=100)

is bigger for higher values of p_f. This indicates a tradeoff between anonymity and performance, as higher p_f implies more intermediate *jondos* in the communication path, and therefore more delay. Regarding the number of members of the crowd, it is clear that the larger the crowd, the higher the value of the effective anonymity set size.

Note that the figures presented correspond to a particular type of attack (namely, the "collaborating *jondos* attack" as described in [8]). The variation of anonymity towards adversaries capable of deploying other attacks may be very different from the results presented in this example. The same applies to the results shown in the next section for the degree of anonymity.

C. Degree of anonymity

The *degree of anonymity* is obtained normalizing the effective anonymity set size with respect to the maximum entropy, H_M. Taking into account that the size of the anonymity set is $N - C$ (the C colluding jondos are not part of the anonymity set), H_M equals:

$$H_M = \log_2 (N - C) \ .$$

According to the formulas presented in Sect. V-C, we compute the *degree of anonymity*, d. Figure 6 represents the degree of anonymity for 20 crowd members, and Fig. 7 for 100 members. As in the figures of the *effective anonymity set size*, the probability of forwarding p_f has been set to 0.5 and 0.75, and the variable in the x axis is the number C of corrupted *jondos*.

We can see in the figures that d decreases with the number of collaborating *jondos* and increases with p_f. The variation of d is very similar for systems with different number of users.

If we compare the results of the two proposed metrics, we can see that while the *effective anonymity set size* presents large variations in $C = 1$ for different values of N (from $\log_2(5) = 2.32$ to $\log_2(100) = 6.64$), the *degree of anonymity* for both crowds systems ($N = 20, 100$) takes values between 0.8 and 0.9 for $p_f = 0.75$, and between 0.6 and 0.7 for $p_f = 0.5$ (evaluated in $C = 1$). Also, the *effective anonymity set size* decreases almost linearly, while the *degree of anonymity* is concave.

The information provided by these metrics can be combined to give a better estimation of the anonymity offered to users. The *effective anonymity set size* gives a quantitative measure of the (un)certainty of the attacker with respect to the identity of a subject, while the *degree of anonymity* indicates the performance of the anonymity system relative to the best it can do for the given number of users. Note that, while the values of the *effective anonymity set size* significantly increase for $N = 100$ with respect to $N = 20$, the *degree of anonymity* slightly decreases for $N = 100$ in comparison with $N = 20$. The explanation to this lies in the fact that, although the *effective anonymity set size* increases due to the increase of potential originators of a communication, the adversary is able to get more information (i.e., reduce his uncertainty with respect to his *a priori* knowledge) from the network with more nodes.

VII. Conclusions

Several solutions for anonymity services have been proposed and implemented in the past. We propose a general model for anonymity systems and present two existing flavors of information theoretic metrics. These metrics provide answers to the research questions formulated in the introduction: they provide a general method to measure anonymity, to compare different systems, to evaluate the effectiveness of attacks on anonymity, to quantify gains and loses in anonymity which take into account the partial or statistical information obtained by an adversary.

With these metrics we can quantify the *effective anonymity set size* and the *degree of anonymity* provided by a system in particular attack circumstances. We have applied the metrics to Crowds, an existing solution for anonymous communication, and discussed the results obtained.

The metrics proposed can be adapted to systems where anonymity can de defined in terms of unlinkability. Anonymous transactions are abstracted as IOIs (*Items Of Interest*); *sender* and *recipient anonymity* can be computed applying the general formulas.

Anonymity metrics provide relevant information on the anonymity of concrete subjects in concrete attack scenarios. In order to know more about the robustness of an anonymity system, we need to make multiple measurements in different scenarios.

The model is based on the probabilities adversaries assign to subjects; finding these probability distributions in real situations is however not always easy.

The question that remains open is the sufficient level of anonymity a system should provide to be privacy enabled. The answer to this question is different for each system, as it depends on the (legal and social) consequences of the breach of privacy in particular scenarios.

References

[1] O. Berthold, and A. Pfitzmann, and R. Standtke, *The disadvantages of free MIX routes and how to overcome them*, H. Federrath (Ed.), Designing Privacy Enhancing Technologies, LNCS 2009, pp. 30-45, 2001.

[2] D. Chaum, *Untraceable electronic mail, return addresses, and digital pseudonyms*, Communications of the ACM 4(2), February 1981.

[3] C. Díaz, and S. Seys, and J. Claessens, and B. Preneel, *Towards measuring anonymity*, Dingledine and Syverson (Eds.), Designing Privacy Enhancing Technologies, LNCS 2482, pp. 54-68, 2002.

[4] C. Díaz, and A. Serjantov, *Generalising Mixes*, Dingledine (Ed.), Designing Privacy Enhancing Technologies, LNCS 2760, pp. 18-31, 2003.

[5] C. Díaz, and B. Preneel, *Reasoning about the Anonymity Provided by Pool Mixes that Generate Dummy Traffic*, Fridrich (Ed.), Information Hiding, LNCS 3200, pp. 309-325, 2004.

[6] C. Díaz, L. Sassaman, and E. Dewitte, *Comparison between two practical mix designs*, ESORICS: 9th European Symposium on Research in Computer Security. LNCS 3193, pp. 141-159, 2004.

[7] A. Pfitzmann, and M. Hansen, *Anonymity, Unobservability, and Pseudonymity: A Proposal for Terminology*, H. Federrath (Ed.), Designing Privacy Enhancing Technologies, LNCS 2009, pp. 1-9, 2000.

[8] M. Reiter, and A. Rubin, *Crowds: Anonymity for Web Transactions*, ACM Transactions on Information and System Security, pp. 66-92, 1998.

[9] A. Serjantov, and G. Danezis, *Towards an Information Theoretic Metric for Anonymity*, Dingledine and Syverson (Eds.), Designing Privacy Enhancing Technologies, LNCS 2482, pp. 41-53, 2002.

[10] Claude E. Shannon, *A Mathematical Theory of Communication*, The Bell System Technical Journal, volume 27:379–423, pp. 623–656, 1948.