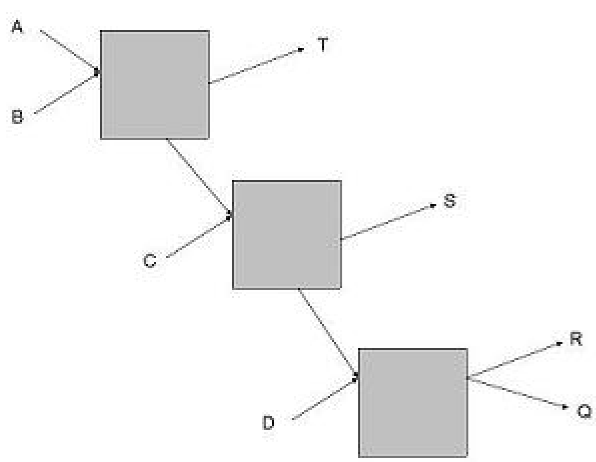
Degree of anonymity

In anonymity networks (e.g., Tor, Crowds, Mixmaster, I2P, etc.), it is important to be able to measure quantitatively the guarantee that is given to the system. The **degree of anonymity d** is a device that was proposed at the 2002 Privacy Enhancing Technology (PET) conference. Two papers put forth the idea of using entropy as the basis for formally measuring anonymity: "Towards an Information Theoretic Metric for Anonymity", and "Towards Measuring Anonymity". The ideas presented are very similar with minor differences in the final definition of d.

Background

Anonymity networks have been developed and many have introduced methods of proving the anonymity guarantees that are possible, originally with simple Chaum Mixes and Pool Mixes the size of the set of users was seen as the security that the system could provide to a user. This had a number of problems; intuitively if the network is international then it is unlikely that a message that contains only Urdu came from the United States, and vice versa. Information like this and via methods like the predecessor attack and intersection attack helps an attacker increase the probability that a user sent the message.

Example With Pool Mixes



As an example consider the network shown above, in here A, B, C and D are users (senders), Q, R, S, and T are servers (receivers), the boxes are mixes, and $\{A, B\} \in T$, $\{A, B, C\} \in S$ and $\{A, B, C, D\} \in Q, R$ where $\in$ denotes the anonymity set. Now as there are pool mixes let the cap on the number of incoming messages to wait before sending be 2; as such if A, B, or C is communicating with R and S receives a message then S knows that it must have come from E (as the links between the mixes can only have 1 message at a time). This is in no way reflected in S's anonymity set, but should be taken into account in the analysis of the network.

Degree of Anonymity

The degree of anonymity takes into account the probability associated with each user, it begins by defining the entropy of the system (here is where the papers differ slightly but only with notation, we will use the notation from [1] (https://en.wikipedia.org/wiki/Degree_of_anonymity#endnote_TMA).):

$H(X) := \sum_{i=1}^{N} \left[p_i \cdot \lg\left(\frac{1}{p_i}\right)\right]$, where $H(X)$ is the entropy of the network, N is the number of nodes in the network, and p_i is the probability associated with node i. Now the maximal entropy of a network occurs when there is uniform probability associated with each node $\left(\frac{1}{N}\right)$ and this yields $H_M := H(X) \leftarrow \lg(N)$. The degree of anonymity (now the papers differ slightly in the definition here, [2] (https://en.wikipedia.org/wiki/Degree_of_anonymity#endnote_TMA) defines a bounded degree where it is compared to H_M and [3] (https://en.wikipedia.org/wiki/Degree_of_anonymity#endnote_TIT) gives an unbounded definition—using the entropy directly, we will consider only the bounded case here) is defined as

$d := 1 - \frac{H_M - H(X)}{H_M} = \frac{H(X)}{H_M}$. Using this anonymity systems can be compared and evaluated using a quantitatively analysis.

Definition of Attacker

These papers also served to give concise definitions of an attacker:

Internal/External

an **internal** attacker controls nodes in the network, whereas an **external** can only compromise communication channels between nodes.

Passive/Active

an **active** attacker can add, remove, and modify any messages, whereas a **passive** attacker can only listen to the messages.

Local/Global

a **local** attacker has access to only part of the network, whereas a **global** can access the entire network.

Example d

In the papers there are a number of example calculations of d; we will walk through some of them here.

Crowds

In Crowds there is a global probability of forwarding (p_f), which is the probability a node will forward the message internally instead of routing it to the final destination. Let there be C corrupt nodes and N total nodes. In Crowds the attacker is internal, passive, and local. Trivially $H_M \leftarrow \lg(N-C)$, and overall the entropy is

$$H(x) \leftarrow \frac{N - p_f \cdot (N-C-1)}{N} \cdot \lg\left[\frac{N}{N - p_f \cdot (N-C-1)}\right] + p_f \cdot \frac{N-C-1}{N} \cdot \lg[N/p_f]$$

, d is this value divided by H_M.[4] (https://en.wikipedia.org/wiki/Degree_of_anonymity#endnote_TMA)

Onion routing

In onion routing, assuming the attacker can exclude a subset of the nodes from the network, the entropy would easily be $H(X) \leftarrow \lg(S)$, where S is the size of the subset of non-excluded nodes. Under an attack model where a node can both globally listen to message passing and is a node on the path this *decreases* to $H(X) \leftarrow \lg(L)$, where L is the length of the onion route (this could be larger or smaller than S), as there is no attempt in onion routing to remove the correlation between the incoming and outgoing messages.

Applications of this metric

In 2004, Diaz, Sassaman, and DeWitte presented an analysis[5] (https://en.wikipedia.org/wiki/Degree_of_anonymity#endnote_CBTPMD) of two anonymous remailers using the Serjantov and Danezis metric, showing one of them to provide zero anonymity under certain realistic conditions.

See also

- Onion routing
- Tor (anonymity network)

- Entropy
- Crowds

References

1. ^ See Towards Measuring Anonymity (http://www.freehaven.net/anonbib/cache/Diaz02.ps.gz) Claudia Diaz and Stefaan Seys and Joris Claessens and Bart Preneel (April 2002). Roger Dingledine and Paul Syverson (ed.). "Towards measuring anonymity" (https://web.archive.org/web/20060710023539/http://www.esat.kuleuven.ac.be/~cdiaz/papers/tmAnon.ps.gz) . *Proceedings of Privacy Enhancing Technologies Workshop (PET 2002)*. Springer-Verlag, LNCS 2482. Archived from the original (http://www.esat.kuleuven.ac.be/~cdiaz/papers/tmAnon.ps.gz) on July 10, 2006. Retrieved 2005-11-10.

2. ^ See Towards an Information Theoretic Metric for Anonymity (https://web.archive.org/web/20040719123728/http://www.cl.cam.ac.uk/~aas23/papers_aas/set.ps) Andrei Serjantov and George Danezis (April 2002). Roger Dingledine and Paul Syverson (ed.). "Towards an Information Theoretic Metric for Anonymity" (https://web.archive.org/web/20040719123728/http://www.cl.cam.ac.uk/~aas23/papers_aas/set.ps) . *Proceedings of Privacy Enhancing Technologies Workshop (PET 2002)*. Springer-Verlag, LNCS 2482. Archived from the original (http://www.cl.cam.ac.uk/~aas23/papers_aas/set.ps) on July 19, 2004. Retrieved 2005-11-10.

3. ^ See Comparison Between Two Practical Mix Designs (http://www.cosic.esat.kuleuven.be/publications/article-98.pdf) Claudia Diaz and Len Sassaman and Evelyn Dewitte (September 2004). Dieter Gollmann (ed.). "Comparison Between Two Practical Mix Designs" (http://www.cosic.esat.kuleuven.be/publications/article-98.pdf) (PDF). *Proceedings of European Symposium on Research in Computer Security (ESORICS 2004)*. Springer-Verlag, LNCS 3193. Retrieved 2008-06-06.