

# QUANTITATIVE MEASURES OF ANONYMITY

By

Matthew J. Edman

A Thesis Submitted to the Graduate  
Faculty of Rensselaer Polytechnic Institute  
in Partial Fulfillment of the  
Requirements for the Degree of  
MASTER OF SCIENCE  
Major Subject: COMPUTER SCIENCE

Approved:

---

Dr. Bülent Yener  
Thesis Adviser

Rensselaer Polytechnic Institute  
Troy, New York

November 2007  
(For Graduation December 2007)

## ABSTRACT

The past two decades have seen a growing interest in methods for anonymous communication on the Internet, both from the academic community and the general public. Several system designs have been proposed in the literature, of which a small number have been implemented and are used by diverse groups, such as journalists, human rights workers, the military, and ordinary citizens, to protect their identities on the Internet.

With multiple anonymity systems and system designs available, it is important to be able to reason about the level of anonymity a particular system provides so that we may compare systems and choose the best one under certain adversarial assumptions. Some previous work has been done to develop quantitative measures for anonymity, but most have only considered anonymity from the perspective of a single user or message in a system. In this thesis, we consider the merits of evaluating anonymity from a system-wide perspective.

We define a new metric for quantifying the degree of anonymity collectively afforded to users of an anonymous communication system. Our approach models the possible relationships between senders and recipients in an anonymity system as a bipartite graph. From the bipartite graph, we derive a metric based on the permanent of that graph's corresponding adjacency matrix to evaluate the amount of information needed by an adversary to reveal the communication pattern as a whole. We show how our model can also be extended to include probabilistic information about relationships between senders and recipients in the system. Finally, we analyze the computational complexity of our metric and discuss lower bounds that can be easily computed.