Quantifying Maximal Loss of Anonymity in Protocols

Han Chen
Queen Mary University of London
School of Electronic Engineering and Computer Science
hanchen@dcs.qmul.ac.uk

Pasquale Malacaria
Queen Mary University of London
School of Electronic Engineering and Computer Science
pm@dcs.qmul.ac.uk

ABSTRACT

There is a natural intuitive match between anonymity and information theory. In particular, the maximal anonymity loss in anonymity protocols can be matched to the information theoretical notion of channel capacity.

However, there is also a significant mismatch between the theories and reality: current theories can only characterize channel capacity based upon certain assumptions of symmetry, which are rarely satisfied in the real world.

This paper aims to resolve this mismatch by appealing to powerful mathematical techniques. A generic methodology using Lagrange multiplier method is proposed to characterize channel capacity in anonymity protocols.

This Lagrangian approach is proved to be able to generalize previous work on the channel capacity of protocols. Further, we present analyses on three well known protocols, namely Dining Cryptographers, Crowds and Onion Routing to demonstrate the application of our methodology.

Categories and Subject Descriptors

C.2.2 [**Network Protocols**]: Protocol verification; G.1.6 [**Optimization**]: Constrained optimization; H.1.1 [**Systems and Information Theory**]: Information theory

General Terms

Security

Keywords

Anonymity, Lagrange multipliers, quantitative analysis

1. INTRODUCTION

Anonymity protocols are playing an increasingly important role in many key fields, such as electronic communication, auction, payment and voting. They are designed to enhance the information privacy of legitimate users while performing certain activities. For example, electronic voting protocols try to protect the confidentiality of individual votes, while anonymous routing protocols (e.g. Onion Routing [24] and Crowds[25]) try to hide the information of "who is communicating with who" when users carry out web activities.

However, these protocols can not completely prevent the loss of anonymity, but instead try to reduce them. For example, if there are a number of adversaries within the onion routing network then a substantial chance exists that the attacker can infer something about the sender. Therefore it is important not only to realize such information leakage, but also to quantify the loss of anonymity.

In this paper, we analyze the anonymity loss over a covert channel, whose input is the "anonymous" events and output is the "observable" events. Then the channel capacity represents exactly the maximal anonymity loss.

An important recent work [4] characterizes maximal loss of anonymity using information theoretical results on binary channels, however its applicability is limited to protocols satisfying certain symmetry properties. Indeed, and this is *the main motivation of this paper*, real world cases rarely satisfy such properties. For example, for anonymity routing protocols the symmetry assumptions will amount to "all nodes having the same probability of being the originator of a message", which is totally inaccurate.

In contrast, we present a methodology using Lagrange multiplier method to solve the channel capacity by maximizing the anonymity loss function. This Lagrangian methodology introduces a generic solution which does not require the symmetry assumptions. It also enables to reflect some interesting relationships in an anonymity system, for example, "Google is 100 times more likely to send a message than server X, Y and 1000 times more likely than any other server". These kind of relationships can be easily represented by *constraints* when solving the Lagrangian equations.

The rest of the paper is organized as follows: Section 2 briefly reviews the Lagrange multiplier method. Section 3 defines the quantification of anonymity loss in protocols. Section 4 introduces our methodology, applying Lagrange method to the channel capacity problem; in Subsection 4.2 we prove that the characterization of channel capacity of symmetric protocols in [4] is a special case of our method. Section 5 and 6 apply the theory to the Dining Cryptographers and Crowds protocols respectively and present the solution to their channel capacity in the asymmetric setting. Section 7 presents a use-case study of Onion Routing, where quantitative measurement is coupled with studies of the impact of the connectivity and path-length settings in the anonymous routing protocol. Section 8 concludes the

ASIACCS'09, March 10-12, 2009, Sydney, NSW, Australia.

paper and outlines our future work.

1.1 Contributions

This paper introduces a methodology for quantifying maximal loss of anonymity in anonymity protocols.

The idea is straightforward but not trivial. We use the classical Lagrange multiplier method to solve a maximization problem that obtains the channel capacity by maximizing the anonymity loss function. With this methodology we preliminarily solved a difficult problem in information theory (pg. 191 from [6]): the channel capacity of asymmetric channels.

Moreover, our approach is both generic and practical. It works for almost any anonymity protocol to enable characterization of the channel capacity between anonymous events and public observations. Especially, in contrast to previous studies it elegantly deals with asymmetric anonymity protocols. Such asymmetric cases are, as argued above, of great practical significance.

Hence the techniques introduced in this paper are also an important contribution to the field: they are not incremental to any previous work in this area and, to the best of our knowledge, this is the first-ever application of this kind of mathematical techniques to the quantitative analysis of protocols.

Further, although the paper focuses on the general methodology, our case studies show insight into the solution of interesting problems in the network security community, e.g. how to optimize protocol parameters (like path length) to enhance the anonymity in an onion routing network.

1.2 Related Works

The mainstream approach to analysis of anonymity protocols is probabilistic, as proposed by M. Reed, P. Syverson and D. Goldschlag [24]. There are many existing works in this topic: Guan et al. [17] measured the probability of a sender being discovered in an anonymous communication system, and quantified the impact from path length, path topology and the number of compromised nodes. Shmatikov and Wang described anonymity in protocols with entropy, and applied Zipfian distributions to estimate the average entropy in a somehow more realistic model [29]. Wright, Adler, Levine and Shields [28] quantitatively analyzed and cross-compared several anonymity protocols according to certain attacks. Recently, J.Feigenmbaum, A. Johnson and P. Syverson in [14] proposed a probabilistic analysis of the anonymity in Onion Routing using a black-box model. There also have been increasing interest on the problem of statistical attacks. Danezis, Diaz and Troncoso [10] analyze a two-sided statistical disclosure attack considering peering and timing information of both initial messages and replies. Pashalidis and Meyer [23] show an attack that links transaction history with pseudonym in pseudonym systems.

Theoretically, a probabilistic approach would be able to work out an expectation of anonymity in certain models; in comparison, this paper is toward finding a general methodology to quantify the maximal loss that may occur, which is an even more important problem but has not yet been satisfactorily solved.

The closest work to ours is by Chatzikokolakis, Palamidessi and Panangaden [4]. Although the work is inspiring and shares the same background with this paper, their methodology only works for "symmetric protocols" and in this sense is a particular case of our work. Further discussions comparing this work with ours will be presented in Subsection 3.3 and 4.2.

Other recent information theoretical approaches to anonymity analysis include [12] and [27]. Also, Shmatikov applied a probabilistic model checking tool Prism to quantitatively analyze anonymity systems [29]. Another interesting work is the one by Franz, Meyer and Pashalidis [15] who analyze anonymity leakage cause by particular "hints" that an adversary may obtain from the context. However, this work is not general and again assumes a uniform distribution on the anonymous events.

The use of conditional mutual information in the context of information leakage has been pioneered by Gray, Denning, McLean and Millen [16, 11, 9, 19, 20]. More recent works using conditional mutual information to measure information leakage have been by Clark, Hunt, Malacaria, Boreale and Chen [7, 8, 22, 2]. These works support this paper from the theory aspect.

In previous literature there was also a general method for solving the problem of channel capacity, namely the Blahut-Arimoto iterative method [6]. While our technique provides a solution via a system of equations, the Blahut-Arimoto method iteratively searches for an approximate solution.

Lagrange multipliers are used by Malacaria and Chen [21] to compute the maximum leakage of deterministic programs. Although sharing similar techniques in between, there is also a significant technical novelty in this paper to cope with the nondeterministic nature of anonymity protocols.

2. LAGRANGE METHOD

We will illustrate the use of the Lagrange method by a simple example below. For formal definitions and a tutorial, we refer the reader to the literature [6, 18].

2.1 A simple example

Suppose we want to maximize the following function:

$$10 - (x-5)^2 - (y-3)^2$$

It is easy to see that the maximum is achieved by $x = 5, y = 3$.

Now a constraint $x+y = 1$ is added to the above problem. Then the above solution is no longer correct. The Lagrange multiplier method combines the original function with the constraint together in a new function F

$$10 - (x-5)^2 - (y-3)^2 + \lambda(x+y-1)$$

where λ is a number which indicates the weight associated with the constraint, for example ignoring the constraint is equivalent to setting $\lambda = 0$.

The term λ is the *Lagrange multiplier* and the Lagrange technique consists in finding the maximum of the function F by differentiating on x, y *and* λ.

In this example the derivatives generate the equations:

$$-2x + 10 + \lambda = 0, \quad -2y + 6 + \lambda = 0, \quad x + y - 1 = 0$$

The first two equations imply $x = y + 2$ and by replacing this in the last equation we get

$$y + 2 + y = 1, \ \textit{i.e.}\ y = -\frac{1}{2}$$

It is then easy to derive the values for the other variables i.e.

$$x = \frac{3}{2},\ \lambda = -7$$

Now the values $x = \frac{3}{2}, y = -\frac{1}{2}$ do satisfy the constraint. They are also the values that maximize the original function

$$10 - (x-5)^2 - (y-3)^2$$

for all values satisfying the constraint. The function evaluated on this point has value -14.5. If we take other values satisfying $x+y=1$ we can only get lower results, e.g. $0.5, 0.5$ results in -16.5 and $1, 0$ results in -15.

2.2 Lagrange Theorem

In a general setting let $L(x, \lambda)$ be the Lagrangian of a function f subject to a family of constraints $C_{1 \le i \le m}$ (where $C_i \equiv g_i(x) = b_i$), i.e.

$$L(x, \lambda) = f(x) + \sum_{1 \le i \le n} \lambda_i (g_i(x) - b_i)$$

The basic result justifying Lagrange multipliers is the following theorem:

THEOREM 2.1. *Assume the vector $x^* = (x_1^*, \ldots, x_n^*)$ maximizes (or minimizes) the function $f(x)$ subject to the constraints $(g_i(x) = b_i)_{1 \le i \le m}$. Then either*

1. *the vectors $(\nabla g_i(x^*))_{1 \le i \le m}$ are linearly dependent, or*
2. *there exists a vector $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)$ such that $\nabla L(\lambda^*, x^*) = 0$ i.e.*

$$(\frac{\delta L}{\delta x_i}(x^*, \lambda^*) = 0)_{1 \le i \le n}, \quad (\frac{\delta L}{\delta \lambda_i}(x^*, \lambda^*) = 0)_{1 \le i \le m},$$

where ∇ is the gradient.

The reverse implication of the theorem is valid when some properties are satisfied. Roughly speaking a maximum is obtained when f is concave and a minimum when f is convex.

The previous example is obtained by the following instantiations:

$$f(x_1, x_2) = 10 - (x_1 - 5)^2 - (x_2 - 3)^2, \quad C \equiv x_1 + x_2 = 1$$

In this paper we will assume that the constraints C_i are "statistics" or expectations , i.e. linear expressions in the form of

$$\sum_j x_j f_{j,i} = F_i$$

More generally, non-linear constraints need to satisfy additional properties of concavity (convexity) for the theory to work.

3. ANONYMITY PROTOCOLS

3.1 Modeling Anonymity Protocol

As presented in [4], anonymity protocols can usually be modeled in a probabilistic setting. We consider an anonymity protocol as a triple

$$\langle \mathcal{A}, \mathcal{O}, \phi \rangle$$

where $\mathcal{A}$ is a set of anonymous events, and $\mathcal{O}$ is a set of observations. To introduce probabilities we associate the random variables h for $\mathcal{A}$ and O for $\mathcal{O}$ respectively. Then, ϕ expresses the conditional probability between the two random variables.

	o_1	o_2	$\ldots$	o_n
h_1	$\phi_{1,1}$	$\phi_{2,1}$	$\ldots$	$\phi_{n,1}$
h_2	$\phi_{1,2}$	$\phi_{2,2}$	$\ldots$	$\phi_{n,2}$
$\vdots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
h_m	$\phi_{1,m}$	$\phi_{2,m}$	$\ldots$	$\phi_{n,m}$

Table 1: Protocol matrix

The "secret" in this model is the information of which event in $\mathcal{A}$ (i.e. which input) caused the observed observation in $\mathcal{O}$. We denote members of $\mathcal{A}$ as $h_i \in \mathcal{A}$.

The above triple can be represented by the protocol matrix shown in Table 1. Rows describe elements of $\mathcal{A}$, columns describe elements of $\mathcal{O}$ and the value at position (h_i, o_k) is the conditional probability $\phi_{k,i}$. This is the chance of observing o_k given h_i as input.

3.2 Channel Capacity

The loss of anonymity of a protocol can be defined as the difference in anonymity before and after the observations, i.e. in information theoretical terms this amount to the mutual information[1] between h and O. Formally, it is defined by the well known information theoretical equation:

DEFINITION 3.1. *The anonymity loss of a protocol $\langle \mathcal{A}, \mathcal{O}, \phi \rangle$ is defined as*

$$I(h; O) = H(h) - H(h|O)$$

where $H(h)$ is the uncertainty of anonymous events in $\mathcal{A}$ and $H(h|O)$ is the remaining uncertainty after observing $\mathcal{O}$.

Consider the following simple example: n voters vote for Clinton or Obama.

- Observations: n ballots which are split in c votes for Clinton and o votes for Obama.
- Anonymity loss: how much information about the identity of a voter for Clinton (or for Obama) is revealed by the observations c and o.

If $\mathcal{O}$ is considered as an output set, the triple $\langle \mathcal{A}, \mathcal{O}, \phi \rangle$ can be regarded as a probabilistic channel. Different distributions on h will result in different anonymity losses $I(h; O)$. What people are interested in is the worst case that can happen in the protocol: what is the maximum value of $I(h; O)$? We use the definition for channel capacity to describe it.

DEFINITION 3.2. *The channel capacity of a protocol $\langle \mathcal{A}, \mathcal{O}, \phi \rangle$ is defined as*

$$C = max\ I(h; O)$$

Up until now we assumed that we know nothing about the anonymous events $\mathcal{A}$. However as pointed out in [4] in most cases some information about the anonymous events is allowed to be revealed by the design of the protocol. In this case there is some knowledge R about the anonymous

[1]For information theory background and notation we refer the reader to [6] or for short summary to [22, 21]

events $\mathcal{A}$, for example in a voting protocol that could be the number of votes for a candidate. This observation leads to an extension of the two previous definitions.

DEFINITION 3.3. *The conditional loss of anonymity of a protocol $\langle \mathcal{A}, \mathcal{O}, \phi \rangle$ with given knowledge R is defined as*

$$I(h; O|R) = H(h|R) - H(h|O, R)$$

DEFINITION 3.4. *The channel capacity of a protocol $\langle \mathcal{A}, \mathcal{O}, \phi \rangle$ with given knowledge R is defined as*

$$C = max\ I(h; O|R)$$

In the special case when R is a function of both h and O, conditional loss of anonymity can somehow be simplified to definitions 3.1 and 3.2.

3.3 Previous Result

It has been shown in previous works that matrices with some sort of symmetry allow a nice characterization of channel capacity [4].

A matrix is symmetric if all rows are permutations of each other and all columns are also permutations of each other. A matrix is weakly symmetric if all rows are permutations of each other and the column sums are equal.

A matrix is partially symmetric (or weakly partially symmetric) if some columns are constant (possibly with different values in each column) and the rest of the matrix is symmetric (or weakly symmetric).

The theorem below has been derived for weakly symmetric matrices [4]:

THEOREM 3.5. **Chatzikokolakis, Palamidessi and Panangaden Theorem:** *Given a protocol described by a weakly symmetric matrix, its channel capacity is given by*

$$C = p_s \log \frac{|\mathcal{O}_s|}{p_s} - H(\mathbf{r}_s)$$

where $\mathcal{O}_s$ is the set of symmetric output values, $\mathbf{r}_s$ the symmetric part of a row of the matrix and p_s is the sum of $\mathbf{r}_s$

This bound is achieved by choosing the uniform input distribution which is hence the channel distribution.

In Subsection 4.2 we will show that Theorem 3.5 can be derived from Proposition 4.1.

4. CHANNEL CAPACITY USING LAGRANGE MULTIPLIERS

We are interested in computing the channel capacity as defined in Definition 3.2 and Definition 3.4. In the following subsections, the Lagrange method will be applied to derive the channel capacity as the solution to the maximization problems.

4.1 Channel Capacity for Anonymity Protocols

Firstly, we are going to apply the Lagrange theorem on anonymity loss of protocols $\langle \mathcal{A}, \mathcal{O}, \phi \rangle$ without given knowledge R. We are hence interested in channel capacity:

$$\max\ I(h; O)$$

Formally, we want to maximize the function

$$f(h) = I(h; O)$$

Convention:

Each anonymous event $h_i \in \mathcal{A}$ is associated to an observation with a given probability $\mu(h_i)$. To ease the exposition we will use h_i both for the event h_i and for its probability $\mu(h_i)$, similarly for o_k. The context will disambiguate what meaning is intended. $\hat{O}_i$ will denote the sets of observations associated to h_i, i.e.

$$\hat{O}_i = \{o_s | \phi_{s,i} \neq 0\}$$

We will use the following properties of $\phi_{k,i}$

$$\sum_i h_i \phi_{k,i} = o_k; \quad \sum_k \phi_{k,i} = 1$$

Notice that

$$\sum_k o_k = \sum_k (\sum_i h_i \phi_{k,i}) = \sum_i h_i (\sum_k \phi_{k,i}) = \sum_i h_i = 1$$

Constraints present the settings of anonymous events $\mathcal{A}$ in protocols. Here we use $\mathcal{C}$ as a set of constraints. Since we are considering a probability distribution a constraint always assumed to be present is $C_0 \equiv \sum h_i = 1$. In the real world, this constraint is not always sufficient. For instance, in the voting example, the people from some voting area might have a higher probability to vote for Clinton than others. Additional constraints can be introduced to specify these conditions, represented by relationship among h_is. Formally, as presented in the last paragraph of Section 2, a constraint $(\mathcal{C}_k)_{k \in K}$ associated to h_i can be decribed as

$$\sum_k f_{i,k} h_i - F_k = 0$$

As described in Section 2.2, the Lagrange method is used to solve this optimization problems with constraints. Following Theorem 2.2, the theorem below solves the distribution of A which will achieve the channel capacity with the constraints. All computations and proofs in this section are omitted because of space limitation.

THEOREM 4.1. *The probabilities h_i maximizing $I(h; O)$ subject to the family of constraint $(\mathcal{C}_k)_{k \in K}$ are given by solving in h_i the equations*

$$\sum_{o_s \in \hat{O}_i} \phi_{s,i} \ln(\frac{\phi_{s,i}}{o_s}) - 1 + \sum_k \lambda_k f_{i,k} \quad = \quad 0$$

and the constraints $(\mathcal{C}_k)_{k \in K}$.

Using this result the channel capacity in this probabilistic channel can be computed.

PROPOSITION 4.2. *The channel capacity is given by*

$$\sum_i h_i (1 - \sum_k \lambda_k f_{i,k}) d$$

where the h_i's are given by theorem 4.1. Moreover, in the case of the single constraint $\sum_i h_i = 1$ the above can be simplified to

$$d(1 - \lambda_0)$$

where $d = \frac{1}{\ln 2}$.

The Theorem 4.1 and Proposition 4.2 can not only be applied in the protocols but also in general probabilistic channels. The following example applies them to the solution of a classical channel capacity problem.

Example: binary symmetric channel

Consider the classic binary symmetric channel (p. 186 [6]) where there are two values for the secret 0,1 and two possible observations 0,1; the probability of the secret being equal to the observation is $1-p$ while the probability of the secret being different from the value observed is p:

$$\phi_{0,0} = \phi_{1,1} = 1-p$$

$$\phi_{0,1} = \phi_{1,0} = p$$

Using $\sum_i h_i \phi_{k,i} = o_k$ we can get

$$o_0 = (1-p)h_0 + ph_1 \quad o_1 = ph_0 + (1-p)h_1$$

Then using Theorem 4.1 we have the equation system:

$$-(1-p)\ln(\frac{o_0}{1-p}) - p\ln(\frac{o_1}{p}) - 1 + \lambda_0 = 0$$
$$-p\ln(\frac{o_0}{p}) - (1-p)\ln(\frac{o_1}{1-p}) - 1 + \lambda_0 = 0$$

By solving it we end up with

$$h_0 = h_1 = \frac{1}{2} \quad \lambda_0 = \ln(\frac{1}{2}) - p\ln(p) - (1-p)\ln(1-p) + 1$$

The channel capacity is then

$$d(1-\lambda_0) = 1 - H(p)$$

which coincide with the classical results on binary symmetric channels [6].

Further, when there is some given knowledge R, then the channel capacity becomes:

$$\max\ I(h; O|R)$$

To solve it, we start by extending $\phi_{k,i}$ to $\phi_{k,i,j}$ to describe the conditional probability of observing o_k given h_i and R_j. Formally,

$$(h_i, R_j, o_k) = (h_i, R_j)\phi_{k,i,j}$$

From the equation above we have

$$\sum_{i,j}(h_i, R_j)\phi_{k,i,j} = o_k; \quad \sum_k \phi_{k,i,j} = 1$$

Notice that

$$\begin{aligned}\sum_k o_k &= \sum_k(\sum_{i,j}(h_i, R_j)\phi_{k,i,j}) \\ &= \sum_{i,j}(h_i, R_j)(\sum_k \phi_{k,i,j}) \\ &= \sum_{i,j}(h_i, R_j) \\ &= 1\end{aligned}$$

Then we use the Lagrange method to figure out the maximum value for $I(h; O|R)$ with the set of constraints $\mathcal{C}$. Here a constraint $(\mathcal{C}_k)_{k\in K}$ associated to (h_i, R_j) can be formally expressed as

$$\sum_k f_{i,j,k}(h_i, R_j) - F_k = 0$$

The Lagrange function becomes:

$$L((h_i, R_j)) = I(h; O|R) + \lambda_k(\sum_k f_{i,j,k}(h_i, R_j) - F_k)$$

The difference to the previous computation is that when we do derivations, we are doing them on the pair (h_i, R_j) instead of the single variable h_i. This is because O is associated with h and R. The concluding theorem is shown below.

THEOREM 4.3. *The probabilities (h_i, R_j) resulting in maximum value of $I(h; O|R)$ subject to the family of constraint $(C_k)_{k\in K}$ are given by solving in (h_i, R_j) the equations*

$$\sum_{o_s \in \hat{O}_{i,j}} \phi_{s,i,j} \ln(\frac{\phi_{s,i,j}}{(o_s|R_j)}) - 1 + \sum_k \lambda_k f_{i,j,k} \quad = \quad 0$$

Then using the probabilities (h_i, R_j) we can work out the channel capacity.

PROPOSITION 4.4. *The channel capacity is given by*

$$\sum_{i,j,k}(h_i, R_j)(1 - \sum_k \lambda_k f_{i,j,k})d$$

where (h_i, R_j)'s are given by theorem 4.3.

4.2 Deriving Theorem 3.5

In this section we are going to show that Theorem 3.5 from [4] is a special case of our Theorem 4.1.

By Proposition 4.1 the probabilities are given by solving h_i in the equations

$$\sum_{o_s \in \hat{O}_i} \phi_{s,i} \ln(\frac{\phi_{s,i}}{o_s}) - 1 + \sum_k \lambda_k f_{i,k} \quad = \quad 0 \qquad (1)$$

In our setting, a weakly symmetric matrix means that there exists a subset of indices K such that given any $k \in K$, for all i, j, $\phi_{k,i} = \phi_{k,j}$. This set is denoted by $\mathcal{O}_n$ in [4]. For all other indices $s \notin K$ we have for all i, j, $(\phi_{s,i})_{s\notin K}$ is a permutation (with no 0 element) of $(\phi_{s,j})_{s\notin K}$: these are the "symmetric output value". To use the same notations as [4], we write r_s for $(\phi_{s,i})_{s\notin K}$ and p_s for $\sum_{s\notin K} \phi_{s,i}$. Also the above conditions imply that for all i, j $\hat{O}_i = \hat{O}_j$. We denote this (unique) set as $\hat{O}$.

As in [4], assuming that there are not additional constraints apart from $\sum_i h_i = 1$ then equation (1) becomes

$$\sum_{o_s \in \hat{O}} \phi_{s,i} \ln(\frac{\phi_{s,i}}{o_s}) - 1 + \lambda_0 \quad = \quad 0 \qquad (2)$$

Using the fact that

$$s \in K \Rightarrow \phi_{s,i} = (o_s|h_i) = o_s$$

It is easy to show that

$$\sum_{o_s \in \hat{O}} \phi_{s,i} \ln(\frac{\phi_{s,i}}{o_s}) - 1 + \lambda_0$$

$$= -\sum_{s\notin K} \phi_{s,i} \ln(o_s) - \ln(2)H(r_s) - 1 + \lambda_0$$

where ln(2) converts log in the entropy formula into the natural logarithm ln. We hence derive the system of equations

$$(\sum_{s\notin K}\phi_{s,i}\ln(o_s) = \ln(2)H(r_s) + 1 - \lambda_0)_{i\in N}$$

Noticing that the right-hand-side is a constant and that for all i, j, $(\phi_{s,i})_{s\notin K}$ is a permutation of $(\phi_{s,j})_{s\notin K}$ we deduce that

$$\forall i, j \notin K, \quad o_i = o_j$$

and since $p_s = \sum_{s\notin K}\phi_{s,j}$ we derive

$$\forall i \notin K, o_i = \frac{p_s}{k},\ k = |\{i \notin K\}|$$

We have hence the equation

$$\sum_{s\notin K}\phi_{s,i}\ln(\frac{k}{p_s}) = \ln(2)H(r_s) + 1 - \lambda_0$$

i.e.

$$p_s\ln(\frac{k}{p_s}) - \ln(2)H(r_s) \quad = \quad 1 - \lambda_0 \qquad (3)$$

Using Proposition 4.2, replacing λ_0 in $d(1-\lambda_0)$ with the left hand side of equation (3) we finally arrive at

$$\frac{1}{\ln(2)}(p_s\ln(\frac{k}{p_s}) - \ln(2)H(r_s)) = p_s\log(\frac{k}{p_s}) - H(r_s)$$

which is Theorem 3.5.

However, if we consider protocols which can be represented by weakly symmetric matrices but the inputs of the protocol has some constraints in addition to $\sum_i h_i = 1$ then Theorem 3.5 is no longer valid.

Recall that when we derive the system of equations

$$(\sum_{s\notin K}\phi_{s,i}\ln(o_s) = \ln(2)H(r_s) + 1 - \sum_k \lambda_k f_{k,i})_{i\in N}$$

The right hand side of the equation is not a constant any more; in particular we cannot derive that

$$\forall i, j \notin K, \quad o_i = o_j$$

Therefore Theorem 3.5 is no longer valid.

In the following three sections we will study three well known anonymity protocols, namely Dining Cryptographers [3], Crowds [25] and Onion Routing [24]. Our methodology will be applied to both symmetric and asymmetric versions of these protocols, in which the results of the symmetric versions are in common with [4]. Furthermore, our methodology also provides accurate results when the symmetry assumption is not satisfied. We will show how the solutions are derived, as well as what this implies to improve the anonymity of these protocols.

5. DINING CRYPTOGRAPHERS

5.1 Protocol description

Three cryptographers are dining on a round table. After the dinner, the master decides who will pay (he or one of the cryptographers), and informs each cryptographer individually about whether he will pay or not. The cryptographers wish to know whether the dinner is payed by one of them or the master, but they also wish to keep the anonymity if one of them is the payer.

Chaum's solution is the following: each cryptographer flips a coin privately and tells the result to the cryptographer on his right. Then each of them compares the coin to his left and his own coin. Each cryptographer will announce N (meaning "disagree") if the two coins are different (head and tail) or Y(meaning "agree") if the two coins are the same (head and head or tail and tail). However, if one of the cryptographers is the payer, he will announce the opposite. If there is an even number of "disagree"s then the Master has paid. Otherwise, the bill has been paid by one of the cryptographer, but the identity of the payer is not revealed to any external observer or the other cryptographers.

h	Coin	O	P
100	000, 111	NYY	$p^3+(1-p)^3$
(h_1)	001, 110	YYN	$p^2(1-p)+(1-p)^2p$
	010, 101	NNN	$p^2(1-p)+(1-p)^2p$
	011, 100	YNY	$p^2(1-p)+(1-p)^2p$
010	000, 111	NYY	$p^2(1-p)+(1-p)^2p$
(h_2)	001, 110	YYN	$p^2(1-p)+(1-p)^2p$
	010, 101	NNN	$p^2(1-p)+(1-p)^2p$
	011, 100	YNY	$p^3+(1-p)^3$
001	000, 111	NYY	$p^2(1-p)+(1-p)^2p$
(h_3)	001, 110	YYN	$p^3+(1-p)^3$
	010, 101	NNN	$p^2(1-p)+(1-p)^2p$
	011, 100	YNY	$p^2(1-p)+(1-p)^2p$
000	000, 111	YYY	$p^3+(1-p)^3$
(Master)	001, 110	NYN	$p^2(1-p)+(1-p)^2p$
	010, 101	YNN	$p^2(1-p)+(1-p)^2p$
	011, 100	NNY	$p^2(1-p)+(1-p)^2p$

Table 2: The Dining Cryptographers protocol

5.2 Anonymity: symmetric case

We denote the three cryptographers sitting around the table as A, B, C. Then the table can be regarded as a ring:
A→ B → C → A

The coin takes value from $\{0, 1\}$; we write p for the probability of the coin being 0 and $1-p$ for 1. Then the protocol is summarized in Table 2.

The first column represents the master's choice. "000" method the master pays the bill while "100, 010, 001" means one of the cryptographer pays. The position of "1" represents the cryptographer who pays the bill. It can be seen from the table that the output set for "$\{100, 010, 001\}$" (i.e. one of the cryptographers pays the bill) is "{NNN, YNY, NYY, YYN}". The output set for "000" (i.e. master pays the bill) is "{YYY, NYN, YNN, NNY}". By observing the outputs one can infer whether the master or the cryptographer pays the bill because the number of "N" is even in the master's output set. Furthermore we define the secret as "which cryptographer pays the bill", i.e. $\{100, 010, 001\}$ and denote it as h_1, h_2, h_3.

5.2.1 Lagrange method

The conditional probabilities ϕ can be written as the weakly symmetric matrix shown in table 3, where

$$a = p^3 + (1-p)^3, \quad b = p^2(1-p) + (1-p)^2p$$

All anonymous events may generate the same observa-

	o_{NYY}	o_{YYN}	o_{NNN}	o_{YNY}
h_1	a	b	b	b
h_2	b	b	b	a
h_3	b	a	b	b

Table 3: Probabilities for Dining Cryptographers

tions, i.e.

$$\hat{O_1} = \hat{O_2} = \hat{O_3} = \{\mathrm{NYY},\ \mathrm{YYN},\ \mathrm{NNN},\ \mathrm{YNY}\}$$

This implies the following probabilities for each observation:

$$\begin{aligned} o_{\mathrm{NYY}} &= \{ah_1 + bh_2 + bh_3\},\ o_{\mathrm{YYN}} = \{bh_1 + bh_2 + ah_3\} \\ o_{\mathrm{NNN}} &= \{bh_1 + bh_2 + bh_3\},\ o_{\mathrm{YNY}} = \{bh_1 + ah_2 + bh_3\} \end{aligned}$$

From this and from Theorem 4.1 we deduce that the channel capacity is given by solving the following equations:

$$\begin{aligned} -a\ln(\tfrac{o_{\mathrm{NYY}}}{a}) - b\ln(\tfrac{o_{\mathrm{YYN}}}{b}) - B - b\ln(\tfrac{o_{\mathrm{YNY}}}{b}) - 1 + \lambda_0 = 0 \\ -b\ln(\tfrac{o_{\mathrm{NYY}}}{b}) - b\ln(\tfrac{o_{\mathrm{YYN}}}{b}) - B - a\ln(\tfrac{o_{\mathrm{YNY}}}{a}) - 1 + \lambda_0 = 0 \\ -b\ln(\tfrac{o_{\mathrm{NYY}}}{b}) - a\ln(\tfrac{o_{\mathrm{YYN}}}{a}) - B - b\ln(\tfrac{o_{\mathrm{YNY}}}{b}) - 1 + \lambda_0 = 0 \end{aligned}$$

where the term $B = -b\ln(\frac{o_{\mathrm{NNN}}}{b}) = -b\ln(\frac{bh_1+bh_2+bh_3}{b}) = 0$ (because $\frac{bh_1+bh_2+bh_3}{b} = 1$) can be eliminated. There is only one λ-term in these equations which is λ_0, and there is only one constraint considered, which is $h_1 + h_2 + h_3 = 1$.

Let us start with an example where the protocol provides perfect anonymity. This is the case if the coin-toss is fair; i.e. $p = \frac{1}{2}$ and therefore $a = b = \frac{1}{4}$. As result of that, the three equations reduce to one:

$$\ln(h_1 + h_2 + h_3) - 1 + \lambda_0 = 0$$

and because $h_1 + h_2 + h_3 = 1$, we get $\lambda_0 = 1$. Now we have the means to calculate the channel capacity by Proposition 4.2. By plugging in the values of λ_0 and h_i we conclude that the channel capacity is 0. Hence there is no loss of anonymity.

For the extreme cases, i.e. when "$p = 0$" and "$p = 1$", this results in $a = 1$ and $b = 0$. The three equations above reduce to

$$\ln(h_i) - 1 + \lambda_0 = 0$$

This system has only one solution, namely

$$h_1 = h_2 = h_3 = \frac{1}{3}$$

which results in a channel distribution of $\log 3$ bits, i.e. the identity of the payer is revealed.

To generalize, with given p, the channel distribution can be solved as above by the only constraint of $\sum_i h_i = 1$, and the channel capacity of the Dining Cryptographer protocol can be calculated by

$$(1-p+p^2)\log 3 - (1-3p+3p^2)\log(\frac{1-p+p^2}{1-3p+3p^2}) - 2(p-p^2)\log(\frac{1-p+p^2}{p-p^2})$$

5.3 Anonymity: asymmetric case

Suppose we now add additional constraints for the distribution of the secret. For example, if the master is ten times more likely to choose the first cryptographer than the second, the information can be represented by a constraint:

$$h_1 = 10h_2$$

using this constraint and from Theorem 4.1 we get the following equations:

$$\begin{aligned} -a\ln(\tfrac{o_{NYY}}{a}) - b\ln(\tfrac{o_{YYN}}{b}) - b\ln(\tfrac{o_{YNY}}{b}) - 1 + \lambda_0 + \lambda_1 = 0 \\ -b\ln(\tfrac{o_{NYY}}{b}) - b\ln(\tfrac{o_{YYN}}{b}) - a\ln(\tfrac{o_{YNY}}{a}) - 1 + \lambda_0 - 10\lambda_1 = 0 \\ -b\ln(\tfrac{o_{NYY}}{b}) - a\ln(\tfrac{o_{YYN}}{a}) - b\ln(\tfrac{o_{YNY}}{b}) - 1 + \lambda_0 = 0 \end{aligned}$$

Using the constraints

$$h_1 = 10h_2, \quad \sum_i h_i = 1$$

this system of equations is simplified to:

$$\begin{aligned} -(1-3b)\ln(A_1) - b\ln(A_2) - b\ln(A_3) = B + 1 - \lambda_0 - \lambda_1 \\ -b\ln(A_1) - b\ln(A_2) - (1-3b)\ln(A_3) = B + 1 - \lambda_0 + 10\lambda_1 \\ -b\ln(A_1) - (1-3b)\ln(A_2) - b\ln(A_3) = B + 1 - \lambda_0 \end{aligned}$$

where $A_1 = 10h_2 - 40bh_2 + b$; $A_2 = 1 - 3b - 11h_2 + 44bh_2$; $A_3 = h_2 - 4bh_2 + b$; $B = -(1-3b)\ln(1-3b) - 2b\ln b$.

These equations only include three unknown variables λ_0, λ_1 and h_2. (b is associated with p which is given.)

By solving these equations the channel capacity is derived using Proposition 4.2.

The channel capacities for Dining Cryptographers in the symmetric (unconstrained) case, and in the case with the additional constraint $h_1 = 10h_2$ are plotted in Figure 1, as a function of p.

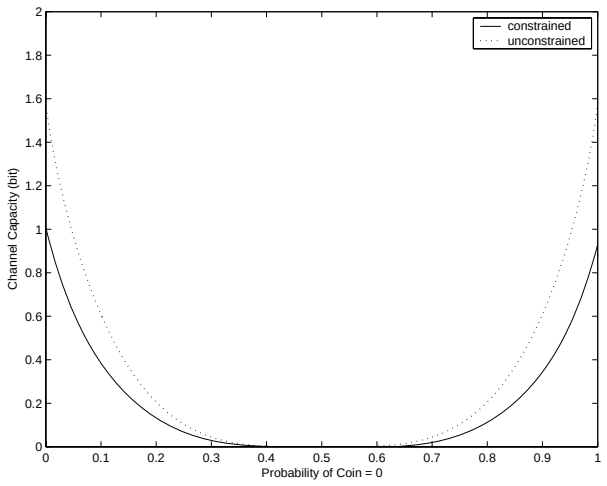

Figure 1: Dining Cryptographers: channel capacity

For both versions of the protocol, when p is 0 or 1 the channel capacity is equal to the entropy of the secret. When the coin becomes fairer, the channel capacity decreases. As p increases from 0 to 0.5, the channel capacity decreases to 0. However, the constraint $h_1 = 10h_2$ implies there is less uncertainty in the initial secret, hence the channel capacity when $p = 0$ is lower compared to the unconstrained case.

6. CROWDS

6.1 Protocol description

The Crowds protocol by Reiter and Rubin [25] enables anonymous Web browsing for end users. The main idea is to hide each user's identity by routing them randomly within a group of similar users.

The routing paths are set up using the following protocol:

- The sender selects a crowd member at random (possibly itself), and forwards the message to it, encrypted by the corresponding pairwise key.
- The selected member, which now acts as router, flips a coin. With probability $1 - p_f$, it delivers the message directly to the destination. With probability p_f, it selects a crowd member at random (possibly itself) as the next router in the path, and forwards the message to it, re-encrypted with the appropriate pairwise key. The next router then repeats this step.

Theoretically, even if a local eavesdropper or a corrupt group member observes a message sent by a particular user, it can not be sure whether the user is the actual sender, or is routing another user's message. Previous result states that, if the crowd contains n members, of which c are corrupt, the minimum value of p_f is required to satisfy the following condition to guarantee the probable innocence of the real sender on any single path [4] (i.e., the probability that the real sender appears on the path immediately before a corrupt member is less than 0.5):

$$n \geq \frac{p_f}{p_f - \frac{1}{2}}(c+1)$$

6.2 Anonymity: symmetric case

Suppose there are n normal users and c corrupted nodes in a network. The attacker is interested in finding the identity of the sender, which is the secret; an observation O is the node being observed (by a corrupted node or the server) to deliver the message. The observations and their probabilities are given in Table 4 where we use numbers from 0 to $n-1$ to identify the normal users. From Table 4, when the secret is h_i, the probability of observation of h_i is $(1-p_f)+p_f\frac{c+1}{n+c}$, in which $(1-p_f)$ comes from the server and $p_f\frac{c+1}{n+c}$ comes from corrupt nodes. This is because h_i has the probability $\frac{c+1}{n+c}$ to choose the corrupt node to forward the request to; the probability of other observations $h_j (j \neq i)$ is $p_f\frac{1}{n+c}$ because they have the same probability to be observed in the routing.

6.2.1 Lagrange method

Now we consider to solve the general case using Lagrange method.

By definition ϕ is:

$$\forall h_i : \ \phi_{i,i} = (1-p_f) + p_f\frac{c+1}{n+c}$$

$$\forall j \neq i : \ \phi_{i,j} = p_f\frac{1}{n+c}$$

Using the relationship between o and ϕ we get:

$$o_i = \sum_{0 \leq j \leq n-1, i \neq j} h_j\phi_{i,j} + h_i\phi_{i,i}$$

h	O	P
0	0	$(1-p_f) + p_f\frac{c+1}{n+c}$
	1	$p_f\frac{1}{n+c}$
	2	$p_f\frac{1}{n+c}$
	3	$p_f\frac{1}{n+c}$
	...	$p_f\frac{1}{n+c}$
	n-2	$p_f\frac{1}{n+c}$
	n-1	$p_f\frac{1}{n+c}$
...	...	...
n-1	0	$p_f\frac{1}{n+c}$
	1	$p_f\frac{1}{n+c}$
	2	$p_f\frac{1}{n+c}$
	3	$p_f\frac{1}{n+c}$
	...	$p_f\frac{1}{n+c}$
	n-2	$p_f\frac{1}{n+c}$
	n-1	$(1-p_f) + p_f\frac{c+1}{n+c}$

Table 4: Crowds: observations and probabilities

This can be rewritten into:

$$o_i = \{\sum_{0 \leq j \leq n-1, i \neq j} (p_f\frac{1}{n+c})h_j + ((1-p_f) + p_f\frac{c+1}{n+c})h_i\}$$

From this and from Theorem 4.1 we have the following system of n equations:

$$-a\ln(\frac{o_0}{a}) - b\ln(\frac{o_1}{b}) - \cdots - b\ln(\frac{o_{n-1}}{b}) - 1 + \lambda_0 = 0$$
$$-b\ln(\frac{o_0}{b}) - a\ln(\frac{o_1}{a}) - \cdots - b\ln(\frac{o_{n-1}}{b}) - 1 + \lambda_0 = 0$$
$$\cdots$$
$$-b\ln(\frac{o_0}{b}) - b\ln(\frac{o_1}{b}) - \cdots - a\ln(\frac{o_{n-1}}{a}) - 1 + \lambda_0 = 0$$

where $a = (1-p_f) + p_f\frac{c+1}{n+c}$; $b = p_f\frac{1}{n+c}$.

This above equations system only admits one solution

$$h_0 = h_1 = h_2 = \cdots = h_{n-1} = \frac{1}{n}$$

The channel capacity is given by

$$I(h;O) = \log n - H(a, \underbrace{b, b, \ldots, b}_{n-1})$$

To compare with the symmetric case in [4], we take their parameters $n = 50$, $c = 10$ to demonstrate our result. Thus from the formula above, the channel capacity is

$$\log 50 + (1 - \frac{49p_f}{60})\log(1 - \frac{49p_f}{60}) + \frac{49p_f}{60}\log\frac{p_f}{60}$$

When the forwarding probability p_f increases, the channel capacity decreases because the attacker has less probability to know who is the sender.

6.3 Anonymity: asymmetric case

In a real world network some users are more active than others. As an example, we study a network with 50 normal users and 10 corrupt ones, where the first four users hold ninety percent of the total probability of sending a message. We further assume that among these four users, the first

user's probability of sending a message is the sum of the second and the third. Because the nonactive users only have very little impact in the whole network, we assume that these only share the remaining ten percent with uniform distribution. Hence the constraints are:

$$h_0 + h_1 + h_2 + h_3 = 0.9,\ h_0 = h_1 + h_2$$

$$h_4 = h_5 = \cdots = h_{49} = \frac{0.1}{46}$$

These constraints imply $\sum_{0 \le i \le 49} h_i = 1$. Then we use equation 4.1 we deduce the following system of 50 equations:

$$\begin{aligned} A_0 + \lambda_0 + \lambda_1 &= 0 \\ A_1 + \lambda_0 - \lambda_1 &= 0 \\ A_2 + \lambda_0 - \lambda_1 &= 0 \\ A_3 + \lambda_0 &= 0 \\ A_k + \lambda_2 &= 0, \quad (4 \le k \le 49) \end{aligned}$$

where

$$A_i = -a \ln W_i - b \sum_{j \neq i} \ln W_j + a \ln a + 49b \ln b - 1$$

and

$$W_r = a h_r + \sum_{s \neq r} b h_s,\ 0 \le r \le 49$$

Because $A_1 = A_2$ we deduce that

$$h_1 = h_2$$

Using

$$h_0 + h_1 + h_2 + h_3 = 0.9, \quad h_0 = h_1 + h_2$$

We can define h_0, h_2, h_3 in terms of h_1 as

$$h_0 = 2h_1, \quad h_2 = h_1, \quad h_3 = 0.9 - 4h_1$$

Also all other values of h can be replaced by a constant, i.e.

$$h_4 = h_5 = \cdots = h_{49} = \frac{0.1}{46}$$

The above system can hence be reduced to a system of 4 equations and 4 unknown variables: $h_1, \lambda_0, \lambda_1, \lambda_2$ which then can be solved for a given p_f using standard numerical analysis methods.

Figure 2 shows the channel capacity of the Crowds protocol in the unconstrained case and constrained cases of 50,1000,10000 users (under the same constraints as in the case of 50 users).

7. ONION ROUTING

Onion Routing [24] is designed to protect data and sender anonymity in communication over a public network such as the Internet. A number of onion routers form an overlay network, in which each onion router is connected to some (if not all) other onion routers. The general idea is, when a client (sender) communicates with a server (receiver), it will first initialize a circuit that comprises of several onion routers. The data will then go through the circuit instead of going directly to the server. We assume that:

1. A circuit can be of any number of nodes as long as no node appears twice.

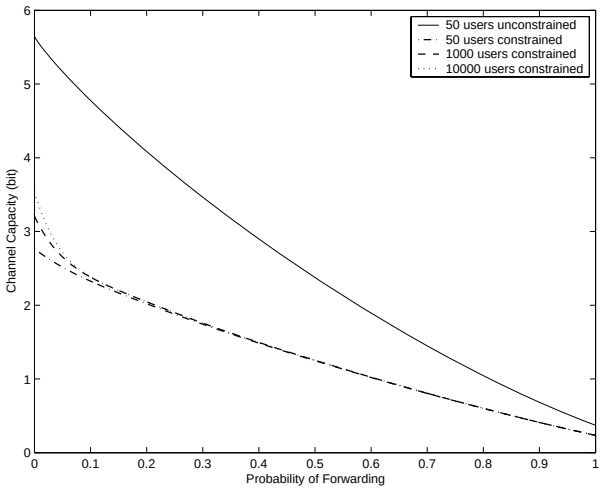

Figure 2: Crowds: channel capacity

2. The client never sends the message to the server directly.

3. Observations by a node include the previous node and the next one.

4. All paths are equally likely.

Ideally, the packet data is encrypted separately for each hop in the circuit so the data confidentiality is protected. The identity of the sender is also partially protected against the server and onion routers, since they do not know whether it is from the sender or another onion router. However, if there are adversaries within the Onion Routing network then there may be a loss for the sender anonymity. [2]

This section will focus on the loss of sender anonymity in this adversary model, and we will show how it can be quantitatively analyzed using the definition of channel capacity. A simple Onion Routing network is used as an example, as shown in Figure 3. The node "R" is the receiver. There are 4 nodes 1,2,3,4 in which either of them can initiate the communication; node 3 is an adversary in the network. We list all the possible paths, observations on the adversary node and the conditional probabilities for the observations in the Table 5.

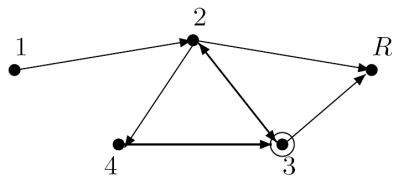

Figure 3: Example of An Onion Routing Network

From the Table 5, we get o using $o_j = \sum_i \phi_{i,j}$:

$$o_{(N,N)} = \frac{1}{3} h_1, \quad o_{(2,R)} = \frac{1}{3} h_1 + \frac{1}{2} h_2$$

[2]An adversary here refers to a compromised node, where the attacker is able to observe which node delivered the packet to it and which node the packet shall then be delivered to.

h	Path	O (in, out)	$\phi_{h_i,O(in,out)}$
$1(h_1)$	$1 \to 2 \to R$	(N, N)	$\frac{1}{3}$
	$1 \to 2 \to 3 \to R$	(2, R)	$\frac{1}{3}$
	$1 \to 2 \to 4 \to 3 \to R$	(4, R)	$\frac{1}{3}$
$2(h_2)$	$2 \to 4 \to 3 \to R$	(4, R)	$\frac{1}{2}$
	$2 \to 3 \to R$	(2, R)	$\frac{1}{2}$
$3(h_3)$	$3 \to 2 \to R$	(N, R)	1
$4(h_4)$	$4 \to 3 \to R$	(4, R)	$\frac{1}{2}$
	$4 \to 3 \to 2 \to R$	(4, 2)	$\frac{1}{2}$

Table 5: Onion Routing: observations and probabilities

$$o_{(4,R)} = \frac{1}{3}h_1 + \frac{1}{2}h_2 + \frac{1}{2}h_4, \quad o_{(N,R)} = h_3, \quad o_{(4,2)} = \frac{1}{2}h_4$$

From o and ϕ and Theorem 4.1 we deduce that the channel capacity is given by solving the following equations:

$$\begin{aligned} -\frac{1}{3}\ln(\frac{o_{(N,N)}}{\frac{1}{3}}) - \frac{1}{3}\ln(\frac{o_{(2,R)}}{\frac{1}{3}}) - \frac{1}{3}\ln(\frac{o_{(4,R)}}{\frac{1}{3}}) - 1 + \lambda_0 = 0 \\ -\frac{1}{2}\ln(\frac{o_{(2,R)}}{\frac{1}{2}}) - \frac{1}{2}\ln(\frac{o_{(4,R)}}{\frac{1}{2}}) - 1 + \lambda_0 = 0 \\ -\ln o_{(N,R)} - 1 + \lambda_0 = 0 \\ -\frac{1}{2}\ln(\frac{o_{(4,2)}}{\frac{1}{2}}) - \frac{1}{2}\ln(\frac{o_{(4,R)}}{\frac{1}{2}}) - 1 + \lambda_0 = 0 \end{aligned}$$

This system has only one solution

$$h_1 = 0.173,\ h_2 = 0.160,\ h_3 = 0.390,\ h_4 = 0.276,\ \lambda_0 = 0.059$$

Using Proposition 4.2, the channel capacity is hence

$$d(1-\lambda_0) = \frac{1}{\ln 2}(1 - 0.0589) = 1.3577 \ \ bits$$

7.1 Anonymity: constrained case

Similar to the analysis of Crowds protocol, we now consider the case when an active user sends out messages more frequently than non-active users. Here we assume h_1 has twice the probability than h_2. Then we have an additional constraint:

$$h_1 = 2h_2$$

with the constraint C_0:

$$h_1 + h_2 + h_3 + h_4 = 1$$

We use Theorem 4.1 to get the following equations:

$$\begin{aligned} -\frac{1}{3}(\ln(\frac{o_{(N,N)}}{\frac{1}{3}}) + \ln(\frac{o_{(2,R)}}{\frac{1}{3}}) + \ln(\frac{o_{(4,R)}}{\frac{1}{3}})) - 1 + \lambda_0 + \lambda_1 = 0 \\ -\frac{1}{2}\ln(\frac{o_{(2,R)}}{\frac{1}{2}}) - \frac{1}{2}\ln(\frac{o_{(4,R)}}{\frac{1}{2}}) - 1 + \lambda_0 - 2\lambda_1 = 0 \\ -\ln o_{(N,R)} - 1 + \lambda_0 = 0 \\ -\frac{1}{2}\ln(\frac{o_{(4,2)}}{\frac{1}{2}}) - \frac{1}{2}\ln(\frac{o_{(4,R)}}{\frac{1}{2}}) - 1 + \lambda_0 = 0 \end{aligned}$$

The system has only one solution

$$h_1 = 0.216, h_2 = 0.108, h_3 = 0.391, h_4 = 0.285,$$

$$\lambda_0 = 0.0615, \lambda_1 = -0.0027$$

h	Path	O (in, out)	$\phi_{h_i,O(in,out)}$
$1(h_1)$	$1 \to 2 \to R$	(N, N)	$\frac{1}{2}$
	$1 \to 2 \to 3 \to R$	(2, R)	$\frac{1}{2}$
$2(h_2)$	$2 \to 3 \to R$	(2, R)	1
$3(h_3)$	$3 \to 2 \to R$	(N, 2)	1
$4(h_4)$	$4 \to 3 \to R$	(4, R)	$\frac{1}{2}$
	$4 \to 3 \to 2 \to R$	(4, 2)	$\frac{1}{2}$

Table 6: The onion network with less connectivity

Using Proposition 4.2 we get the channel capacity(in bits):

$$d(h_1(1-\lambda_0-\lambda_1)+h_2(1-\lambda_0+2\lambda_1)+(h_3+h_4)(1-\lambda_0)) = 1.354$$

We have only one constraint in this case, and from the formula in Theorem 4.1

$$\sum_{o_s \in \hat{O}_i} \phi_{s,i} \ln(\frac{\phi_{s,i}}{o_s}) - 1 + \sum_k \lambda_k f_{i,k} = 0$$

multiple constraints will only affect the last item $\sum_k \lambda_k f_{i,k}$ in the equation system. The complexity is increased linearly by increasing the number of factors λ_k.

7.2 Anonymity: the impact of network connectivity

In an onion network each individual router can allow or disallow connections from other routers. Using the same additional constraint as in the previous case study, we set off to investigate the relationship between the connectivity and the loss of anonymity in an onion network. This can be the basis to understand quantitatively how much connectivity an onion network needs to have to achieve a certain level of anonymity.

In this case, the connection from node 2 to node 4 is removed, as shown in Figure 4. The additional constraint is maintained:

$$h_1 = 2h_2$$

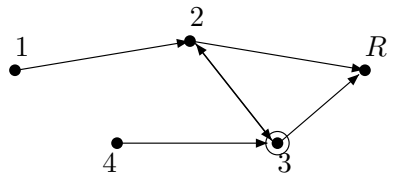

Figure 4: An Onion Network with Less Connectivity

The observable outputs and corresponding probabilities are listed in Table 6. From the table, we can get o using $o_j = \sum_i \phi_{i,j}$ as follows:

$$o_{(N,N)} = \frac{1}{2}h_1, \quad o_{(2,R)} = \frac{1}{2}h_1 + h_2$$

$$o_{(4,R)} = \frac{1}{2}h_4, \quad o_{(N,2)} = h_3, \quad o_{(4,2)} = \frac{1}{2}h_4$$

From o and ϕ and Theorem 4.1 we deduce that the channel distribution of h_i's is given by solving the following equa-

tions:

$$-\frac{1}{2}\ln(\frac{o_{(N,N)}}{\frac{1}{2}}) - \frac{1}{2}\ln(\frac{o_{(2,R)}}{\frac{1}{2}}) - 1 + \lambda_0 + \lambda_1 = 0$$
$$-\ln(\frac{o_{(2,R)}}{1}) - 1 + \lambda_0 - 2\lambda_1 = 0$$
$$-\ln o_{(N,2)} - 1 + \lambda_0 = 0$$
$$-\frac{1}{2}\ln(\frac{o_{(4,R)}}{\frac{1}{2}}) - \frac{1}{2}\ln(\frac{o_{(4,2)}}{\frac{1}{2}}) - 1 + \lambda_0 = 0$$

With the constraints the system has only one solution

$$h_1 = 0.2488, h_2 = 0.1244, h_3 = 0.2834, h_4 = 0.2834$$

$$\lambda_0 = -0.2609, \lambda_1 = 0.1155$$

Using Proposition 4.2 we get the channel capacity in a similar way (in bits)

$$d(h_1(1-\lambda_0-\lambda_1)+h_2(1-\lambda_0+2\lambda_1)+(h_3+h_4)(1-\lambda_0)) = 1.819$$

Notice there is hence a difference of 0.54 bits between 1.891 and 1.354 as in the previous case study, which is due to the reduced connectivity of the second network. As the intuition suggests, if there is better connectivity and generally more paths to choose from, then there will be higher entropy with respect to the observations, and better anonymity can be achieved. This example shows that our techniques can be used to measure the impact from certain parameters of a protocol on the loss of its anonymity.

7.3 Anonymity: Knowing the path length

In this section, we are going to show how the Theorem 4.3 and Proposition 4.4 can be applied to study the impact of tuning the path length in Onion Routing.

In an onion network, suppose the protocol sets the length of the path which is also known to the attacker. In this case, the path length is represented by additional, "low" information in theoretical terms.

For illustration, we use the same case as in Figure 4 without any constraint. Here R_j represents the probability that a path has a length of j. For example, we use R_2 for the path $1 \to 2 \to R$ because the length is 2.

We will show that the channel capacity hereby derived will be a linear combination of R_j, which implies that tuning R_j will effectively change the channel capacity.

In this case, the constraints are only built up from the joint probability of the input:

$$\sum_i (h_i, R_j) = R_j$$

For this example, the observable outputs and corresponding joint probabilities are listed in Table 7. Here the secret is listed in the first column as before; the path lengths are either 2 or 3 which are interpreted as R_2 and R_3 in the second one; the probabilities (o_s, R_j) are (h_i, R_j) because each pair has different output.

From Table 7, each input pair (h_i, R_j) produces different outputs, thus the conditional probabilities ϕ is:

$$\phi_{i,j,k} = 1$$

From o and ϕ and Theorem 4.3 we deduce that the channel capacity is given by solving the following equations:

h	R	Path	O (in, out)	(o_s, R_j)
$1(h_1)$	R_2	$1 \to 2 \to R$	(N, N)	(h_1, R_2)
	R_3	$1 \to 2 \to 3 \to R$	(2, R)	(h_1, R_3)
$2(h_2)$	R_2	$2 \to 3 \to R$	(2, R)	(h_2, R_2)
$3(h_3)$	R_2	$3 \to 2 \to R$	(N, 2)	(h_3, R_2)
$4(h_4)$	R_2	$4 \to 3 \to R$	(4, R)	(h_4, R_2)
	R_3	$4 \to 3 \to 2 \to R$	(4, 2)	(h_4, R_3)

Table 7: The onion network with less connectivity

$$\ln\frac{R_2}{(h_1,R_2)} - 1 + \lambda_0 = 0 \quad \ln\frac{R_3}{(h_1,R_3)} - 1 + \lambda_1 = 0$$
$$\ln\frac{R_2}{(h_2,R_2)} - 1 + \lambda_0 = 0 \quad \ln\frac{R_2}{(h_3,R_2)} - 1 + \lambda_0 = 0$$
$$\ln\frac{R_2}{(h_4,R_2)} - 1 + \lambda_0 = 0 \quad \ln\frac{R_3}{(h_4,R_3)} - 1 + \lambda_1 = 0$$

With the constraints the system admits the solution:

$$h_1 = \frac{1}{4}R_2 + \frac{1}{2}R_3, \quad h_2 = \frac{1}{4}R_2, \quad h_3 = \frac{1}{4}R_2$$

$$h_4 = \frac{1}{4}R_2 + \frac{1}{2}R_3, \quad \lambda_0 = 1 + \ln\frac{1}{2}, \quad \lambda_1 = 1 + \ln\frac{1}{4}$$

Using Proposition 4.4 we get the channel capacity:

$$d\sum_{i,j,k}(h_i, R_j)(1 - \sum_k \lambda_k f_{i,j,k}) = R_2 + 2R_3 \text{ bits}$$

which is a linear combination of R_2 and R_3. For example, if we assume $R_2 = R_3$ then the corresponding channel capacity is

$$R_2 + 2R_3 = 1.5 \text{ bits}$$

In fact, when R_j is introduced into the system of equations, the solution of the system of equations becomes a linear combination involving R_j. The channel capacity (maximal anonymity loss) is then a linear function of R_j. Therefore, ideally we can find an optimum probabilistic distribution of the path length that will result in the best anonymity.

An important implication is that it becomes possible to tune the parameters within the protocol (e.g. path length) to automatically adapt to dynamic properties of the anonymity system (e.g. possible location and distribution of adversaries and level of connectivity). It is beyond the scope of this paper to discuss the specific algorithms used by Onion Routing, however we believe an opportunity clearly exists to improve such protocols in the light of our analysis.

8. CONCLUSION

This paper introduced Lagrange multipliers to analyse the channel capacity in anonymity protocols. The constraints in the Lagrange method are shown to have a practical relevance e.g. in reflecting real properties of networks, in quantifying the design weaknesses in anonymity protocols, but also to specify and analyze the impact of properties like network connectivity.

We believe our work, as a theoretical framework, is accurate, useful and feasible. We applied the methodology to analyze three different protocols. More applications can certainly be found in such areas like communications, voting and auctions.

8.1 Further work

Scalability: The solution of system of equations can be very costly as the number of unknown (constraints) grows. There is a wealth of knowledge in the literature on efficient solutions to systems of equations which we haven't explored so far.

Inequality and Nonlinear Constraints: In this paper we have dealt with constraints expressed as linear equations. We are currently working on the use of nonlinear constraints. Further, to deal with inequalities as constraints we need to use a generalization of the technique known as Karush-Kuhn-Tucker conditions.

Probabilistic analysis: A comparison of the information theoretical and probabilistic analysis of anonymity protocols would be very interesting. For example It would be interesting to compare our analysis with the probabilistic analysis of Onion routing presented in [14].

9. REFERENCES

[1] Mohit Bhargava, Catuscia Palamidessi. Probabilistic Anonymity. CONCUR 2005, LNCS 3653, pp. 171-185, 2005.

[2] Michele Boreale: Quantifying Information Leakage in Process Calculi. ICALP (2) 2006: 119-131

[3] David Chaum: The Dining Cryptographers Problem: Unconditional Sender and Recipient Untraceability. Journal of Cryptology 1 (1): 65-75, 1988.

[4] K. Chatzikokolakis , C. Palamidessi , P. Panangaden. Anonymity Protocols as Noisy Channels, in: Postproceedings of the Symp. on Trustworthy Global Computing, Lecture Notes in Computer Science, Springer, 2006.

[5] Han Chen, Pasquale Malacaria: Quantitative Analysis of Leakage for Multi-threaded Programs. Proc. ACM 2007 workshop on Programming languages and analysis for security.

[6] T.Cover, J. Thomas. Elements of Information Theory. Wiley

[7] David Clark, Sebastian Hunt, Pasquale Malacaria: A static analysis for quantifying information flow in a simple imperative language. Journal of Computer Security, Volume 15, Number 3 / 2007.

[8] David Clark, Sebastian Hunt, Pasquale Malacaria: Quantitative Analysis of the leakage of confidential data. Electronic Notes in Theoretical Computer Science 59, 2002

[9] D. E. Denning: Cyptography and Data Security. Addison-Wesley, 1982.

[10] George Danezis, Claudia Diaz and Carmela Troncoso. Two-Sided Statistical Disclosure Attack. In Proc. PET 2007, LNCS 4476, pp. 30-44, 2007.

[11] D. E. Denning: A lattice model of secure information flow. Communications of the ACM, 19(5), May 1976.

[12] C. Díaz, S.Seys, J Claessens and B. Preneel: Towards measuring anonymity. Proceedings of Privacy Enhancing Technologies Workshop (PET 2002),Springer-Verlag, LNCS 2482.

[13] Matthew Edman, Fikret Sivrikaya, Bulent Yener. A Combinatorial Approach to Measuring Anonymity. In Proc. Intelligence and Security Informatics, 2007 IEEE, 2007, pages 356-363.

[14] J.Feigenmbaum, A. Johnson, P. Syverson. Probabilistic Analysis of Onion Routing in a Black-box Model. In proceedings of WPES'07, ACM 2007.

[15] Matthias Franz, Bernd Meyer and Andreas Pashalidis, Attacking Unlinkability: The Importance of Context. In Proc. PET 2007, LNCS 4776, pp. 1-16, 2007.

[16] James W Gray III: Toward a methematical foundataion for information flow security. Proc. 1991 IEEE Symposium on Security and Privacy. Oakland, California, May 1991.

[17] Yong Guan, Xinwen Fu, Riccardo Bettati, and Wei Zhao: A quantitative analysis of anonymous communications. IEEE Transactions on Reliability, Page 103-115, Volume 53(1), March 2004.

[18] S. Kullback: Information Theory and Statistics. Dover Publications. 1997.

[19] John Mclean: Security models and information flow. Proc. 1990 IEEE Symposium on Security and Privacy. Oakland, California, May 1990.

[20] Jonathan Millen: Covert channel capacity. Proc. 1987 IEEE Symposium on Research in Security and Privacy.

[21] Pasquale Malacaria, Han Chen,: Lagrange Multipliers and Maximum Information Leakage in Different Observational Models. ACM SIGPLAN Third Workshop on Programming Languages and Analysis for Security. June, 2008.

[22] Pasquale Malacaria: Assessing security threats of looping constructs. Proc. ACM Symposium on Principles of Programming Language, 2007.

[23] Andreas Pashalidis and Bernd Meyer Linking Anonymous Transactions: The Consistent View Attack. In Proc. PET 2006, LNCS 4258, pp. 384-392, 2006.

[24] M. Reed, P. Syverson, D. Goldschlag,"Anonymous Connections and Onion Routing", IEEE Symposium on Security and Privacy (1997)

[25] M. Reiter and A. Rubin. Crowds: Anonymity for web transactions. ACM Transactions on Information and System Security (TISSEC), 1(1), pages 66-92. 1998.

[26] C. E. Shannon and W. Weaver: A Mathematical Theory of Communication. Urbana, IL: Univ. of Illinois press, 1963.

[27] A. Serjantov, G.Danezis. Towards an Information Theoretic Metric for Anonymity. Proceedings of Privacy Enhancing Technologies Workshop (PET 2002),Springer-Verlag, LNCS 2482.

[28] M. Wright, M. Adler, B. Levine, and C. Shields. An analysis of the degradation of anonymous protocols. In ISOC Network and Distributed System Security Symposium (NDSS), 2002.

[29] Vitaly Shmatikov. Probabilistic model checking of an anonymity system. Journal of Computer Security, vol 12, 2004.

[30] Vitaly Shmatikov and Ming-hsiu Wang. Measuring Relationship Anonymity in Mix Networks. In proceedings of WPES'06, ACM 2006.