Dynamic Multi–Threshold Metering Schemes

Carlo Blundo, Annalisa De Bonis, Barbara Masucci

Dipartimento di Informatica ed Applicazioni
Università di Salerno, 84081 Baronissi (SA), Italy

E-mail: {carblu, debonis, masucci}@dia.unisa.it

Douglas R. Stinson

Department of Combinatorics and Optimization
University of Waterloo
Waterloo, Ontario, N2L 3G1, Canada

E-mail: dstinson@cacr.math.uwaterloo.ca

Abstract

A *metering scheme* is a method by which an audit agency is able to measure the interaction between clients and servers on the web during a certain number of time frames. Naor and Pinkas [7] considered metering schemes in which any server is able to construct a proof to be sent to the audit agency if and only if it has been visited by at least a number, say h, of clients in a given time frame. In their schemes the parameter h is fixed and is the same for any server and any time frame. This is acceptable whenever there is a long-term relationship between the audit agency and the servers.

In order to measure any number of visits in any granularity we introduce *dynamic multi–threshold metering schemes*, which are metering schemes in which there is a threshold h_j^t associated to any server $\mathcal{S}_j$ for any time frame t. We mainly focus on the efficiency of dynamic multi–threshold metering schemes, providing lower bounds on the size of the information distributed to clients and to servers (this is important otherwise the task of receiving and sending information would burden the clients, that are not interested in the metering process).

Keywords: Metering Schemes, Security, Cryptography, Entropy.

1 Introduction

Most of the revenues of web sites come from advertisement payments. Web advertisers must have a way to measure the exposure of their ads by obtaining usage statistics about web sites which contain their ads. Indeed, the amount of money charged to display ads depends on the number of visits received by the web site. Consequently, advertisers should prevent the web sites from inflating the count of their visits in order to demand more money. Hence, there should be a mechanism which ensures the validity and accuracy of usage measurements against fraud attempts by servers (web sites) and clients (visitors). In a typical scenario there are many servers and clients, and an audit agency whose task is to measure the interaction between the servers and the clients. A system for measuring the amount of service performed by the servers is called *metering scheme*.

Franklin and Malkhi [5] were the first to consider the secure metering problem in a rigorous theoretical approach. They proposed solutions which offer "lightweight security" against corrupted servers which try to inflate the count of their visits. Such solutions cannot be applied if there are strong commercial interest to falsify the metering result, which is the case of many advertising applications on the web.

Consequently Naor and Pinkas [7] proposed metering schemes secure against coalitions of corrupt servers and clients. The main property of their metering schemes is that any server is able to present to the audit agency a short proof for the number of client visits it has received in a given time frame. In their schemes all servers are associated to a threshold h, and are able to compute their proof for a certain time frame if and only if they have been visited by a number of clients larger than or equal to h in that time frame. The schemes proposed by Naor and Pinkas are also efficient: the task for the audit agency in sending information to clients and servers is very simple, as well as the task for the servers in computing their proofs.

In metering schemes considered by Naor and Pinkas [7] the parameter h is fixed and is the same for any server and any time frame. This is acceptable whenever there is a long-term relationship between the audit agency and the servers. In order to measure any number of visits in any granularity we introduce *dynamic multi–threshold metering schemes*, which are metering schemes in which there is a threshold h_j^t associated to any server $\mathcal{S}_j$ for any time frame t.

Dynamic multi–threshold metering schemes involves distributing information to clients and servers. Obviously, such information distribution affects the overall communication complexity. Therefore it is important to construct schemes whose overhead to the overall communication is as small as possible. We mainly focus on the size of information distributed to clients and to servers in dynamic multi–threshold metering schemes. This is important otherwise the task of receiving and sending information would burden the clients, that are not interested in the metering process. We provide lower bounds on the size of the information distributed to clients and on the size of the information given from clients to servers during their visits.

2 The Model

Consider the following scenario: there are n clients, m servers and an audit agency A which is interested in counting the client visits to the servers in τ different time frames. For any $i = 1, \ldots, n$ and $j = 1, \ldots, m$, we denote by $\mathcal{C}_i$ the i-th client and by $\mathcal{S}_j$ the j-th server.

There is an *initialization phase* in which the audit agency A distributes some information to any client over a private channel. For any $i = 1, \ldots, n$, we denote by c_i the information that the audit agency A gives to the client $\mathcal{C}_i$. Moreover, we denote by C_i the set of all values that c_i can assume.

At the *beginning of any time frame* the audit agency A distributes to any server a piece of information which depends on the identity of the server and on the time frame. For any $j = 1, \ldots, m$ and $t = 1, \ldots, \tau$, we denote by s_j^t the information that the audit agency A gives to the server $\mathcal{S}_j$ at the beginning of time frame t. Moreover, we denote by S_j^t the set of all values that s_j^t can assume.

A *regular operation* consists in a client visit to a server during a time frame. During such a visit the client gives to the visited server a piece of information which depends on its private information, on the identity of the server, and on the time frame during which the client visits the server. For any $i = 1, \ldots, n$, $j = 1, \ldots, m$, and $t = 1, \ldots, \tau$, we denote by $c_{i,j}^t$ the information that the client $\mathcal{C}_i$ sends to the server $\mathcal{S}_j$ when visiting it in time frame t. Moreover, we denote by $C_{i,j}^t$ the set of all values that $c_{i,j}^t$ can assume. For any $j = 1, \ldots, m$ and $t = 1, \ldots, \tau$, we denote by $X_{j,(d_j)}^t$ the set of the d_j client visits received by server $\mathcal{S}_j$ in time frame t.

During the *proof computation stage* any server $\mathcal{S}_j$ which has received at least h_j^t visits during time frame t is able to compute its proof for time frame t, as function of the information provided by the h_j^t clients and the information provided by the audit agency A at the beginning of the time frame t. For any $j = 1, \ldots, m$ and $t = 1, \ldots, \tau$, we denote by p_j^t the proof computed by the server $\mathcal{S}_j$ when it has been visited by at least h_j^t distinct clients in time frame t. Moreover, we denote by P_j^t the sets of all values that p_j^t can assume.

During the *proof verification stage* the audit agency A verifies the proofs received by servers and decides on the amount of money to be paid to servers. If the proof received from a server at the end of a time frame is correct, then A pays the server for its services.

A *corrupt* server can be assisted by corrupt clients and other corrupt servers in order to inflate the count of its visits. A corrupt client $\mathcal{C}_i$ can donate to a corrupt server the whole private information received by the audit agency during the initialization phase. We assume that the number of corrupt clients is c, where $1 \leq c < \min_{j=1,\ldots,m} \min_{t=1,\ldots,\tau} h_j^t$. A corrupt server can donate to another corrupt server the private information received from the audit agency at the beginning of any time frame plus the information received from clients in previous time frames and in the actual time frame. For any $i = 1, \ldots, n$ and $t = 1, \ldots, \tau$, we denote by $V_j^{[t]}$ all the information received by a corrupt server $\mathcal{S}_j$ in time frames $1, \ldots, t$. This information includes the sets of client visits received by server $\mathcal{S}_j$ in time frames $1, \ldots, t$. We also define $V_j^{[0]} = \emptyset$, for any corrupt server $\mathcal{S}_j$. We assume that the maximum number of corrupt servers is s, where $1 \leq s \leq m$. For the reader's convenience, the notations used in this section are summarized in the table of Figure 1.

In this paper with a boldface capital letter, say $\mathbf{X}$, we denote a random variable taking value on a set denoted by the corresponding capital letter X according to some probability distribution $\{Pr_{\mathbf{X}}(x)\}_{x \in X}$. The values such a random variable can take are denoted by the corresponding lower letter. Let d an arbitrary positive integer and let $\mathbf{X}_1, \ldots, \mathbf{X}_d$ be d random variables taking values on the sets $X_1, \ldots, X_d$, respectively. For any subset $V = \{i_1, \ldots, i_v\} \subseteq \{1, \ldots, d\}$, with $i_1 \leq \ldots \leq i_v$, we denote with X_V the set $X_{i_1} \times \ldots \times X_{i_v}$ and with $\mathbf{X}_V$ the sequence of random variables $\mathbf{X}_{i_1}, \ldots, \mathbf{X}i_v$. Given a random variable $\mathbf{X}$ we denote with $H(\mathbf{X})$ the Shannon entropy of $\{Pr_{\mathbf{X}}(x)\}_{x \in X}$ (for some basic properties of entropy, consult the Appendix).

We formally define multi–threshold metering schemes by using the entropy approach, as done in [4], [2], and [6]. We use the entropy approach mainly because this leads to a compact and simple description of the schemes and because the entropy approach takes into account all probability distributions on the sets of the proofs computed by the servers.

Definition 2.1 *A dynamic multi–threshold metering scheme* $\mathcal{M} = (n, m, \tau, c, s, \{h_j\}_{j=1,\ldots,m}^{t=1,\ldots,\tau})$ *is a protocol to measure interaction between* n *clients and* m *servers during* τ *time frames in such a way that the following properties are satisfied:*

1. *Any client is able to compute the information needed to visit any server in any time frame:* Formally, it holds that $H(\mathbf{C}_{i,j}^t | \mathbf{C}_i) = 0$ for $i = 1, \ldots, n$, $j = 1, \ldots, m$, and $t = 1, \ldots, \tau$.

2. *Any server* $\mathcal{S}_j$ *which has received* h_j^t *client visits during time frame* t *and the message provided by* A *at the beginning of the time frame* t *can compute its proof for* t*:* Formally, it holds that $H(\mathbf{P}_j^t | \mathbf{X}_{j,(h_j^t)}^t \mathbf{S}_j^t) = 0$, for $j = 1, \ldots, m$ and $t = 1, \ldots, \tau$.

3. *Let us consider a coalition of α corrupt clients $\mathcal{C}_1, \ldots, \mathcal{C}_\alpha$ and β corrupt servers $\mathcal{S}_1, \ldots, \mathcal{S}_\beta$, where $0 \le \alpha \le c < \min_{j=1,\ldots,m} \min_{t=1,\ldots,\tau} h_j^t$ and $1 \le \beta \le s$, and let $B = \{1, \ldots, \beta\}$. Assume that at some time frame t each server $\mathcal{S}_j$ in the coalition has been visited by less than $h_j^t - \alpha$ clients and has received the information by A. Then, the servers in the coalition have no information on their proofs for t:*
   Formally, it holds that $H(\mathbf{P}_B^t | \mathbf{C}_1 \ldots \mathbf{C}_\alpha \mathbf{X}_{1,(d_1)}^t \ldots \mathbf{X}_{\beta,(d_\beta)}^t \mathbf{S}_B^t \mathbf{V}_B^{[t-1]}) = H(\mathbf{P}_B^t)$, where $d_j < h_j^t - \alpha$, for $j = 1, \ldots, \beta$.

Notice that this model is different from the situation considered by Naor and Pinkas [7]. Indeed, they considered metering schemes which were "static" and with "single threshold", i.e., where $h_j^t = h$ for $j = 1, \ldots, m$ and $t = 1, \ldots, \tau$. Moreover, their schemes do not require communication between audit agency and servers at the beginning of any time frame.

3 Lower Bounds on the Size of the Information Distributed to Clients and Servers

Dynamic multi–threshold metering schemes involve distributing information to clients and servers. In a dynamic multi–threshold metering scheme the clients receive a certain amount of information from the audit agency and give part of this information to the servers when visiting them.

We mainly focus on the size of information distributed to clients and to servers in dynamic multi–threshold metering schemes. This is important otherwise the task of receiving and sending information would burden the clients, that are not interested in the metering process. In this section we provide lower bounds on the size of the information distributed to clients and servers in dynamic multi–threshold metering schemes.
In order to prove our results we will resort to the two following technical lemmas.

Lemma 3.1 *Let $\mathbf{A}$ and $\mathbf{E}$ be two random variables such that $H(\mathbf{A}|\mathbf{E}) = 0$. Then, for any two random variables $\mathbf{F}$ and $\mathbf{G}$, it holds that*

$$H(\mathbf{G}|\mathbf{AEF}) = H(\mathbf{G}|\mathbf{EF}).$$

Proof. Consider the mutual information $I(\mathbf{A}; \mathbf{G}|\mathbf{EF})$. ¿From (10) of Appendix it holds that

$$H(\mathbf{A}|\mathbf{EF}) - H(\mathbf{A}|\mathbf{EFG}) = H(\mathbf{G}|\mathbf{EF}) - H(\mathbf{G}|\mathbf{AEF}).$$

¿From (11) of Appendix we have that $H(\mathbf{A}|\mathbf{EFG}) \le H(\mathbf{A}|\mathbf{EF}) \le H(\mathbf{A}|\mathbf{E})$. Since $H(\mathbf{A}|\mathbf{E}) = 0$, it follows that

$$H(\mathbf{G}|\mathbf{AEF}) = H(\mathbf{G}|\mathbf{EF}).$$

□

Lemma 3.2 *Let $\mathbf{E}$, $\mathbf{F}$, and $\mathbf{G}$ be three random variables such that $H(\mathbf{G}|\mathbf{EF}) = 0$ and $H(\mathbf{G}|\mathbf{E}) = H(\mathbf{G})$. Then, it holds that*

$$H(\mathbf{F}|\mathbf{E}) = H(\mathbf{G}) + H(\mathbf{F}|\mathbf{EG}).$$

Proof. Consider the mutual information $I(\mathbf{F}; \mathbf{G}|\mathbf{E})$. ¿From (10) of Appendix it holds that

$$H(\mathbf{F}|\mathbf{E}) - H(\mathbf{F}|\mathbf{EG}) = H(\mathbf{G}|\mathbf{E}) - H(\mathbf{G}|\mathbf{EF}).$$

Since $H(\mathbf{G}|\mathbf{EF}) = 0$ and $H(\mathbf{G}|\mathbf{E}) = H(\mathbf{G})$, then it follows that $H(\mathbf{F}|\mathbf{E}) = H(\mathbf{G}) + H(\mathbf{F}|\mathbf{EG})$.
□

The next lemma immediately follows from Definition 2.1. Recall that for any sets of clients and servers indices $A = \{1, \ldots, \alpha\} \subseteq \{1, \ldots, n\}$ and $B = \{1, \ldots, \beta\} \subseteq \{1, \ldots, m\}$, respectively, we denote by $c^t_{A,B}$ the information given by clients $\mathcal{C}_1, \ldots, \mathcal{C}_\alpha$ to servers $\mathcal{S}_1, \ldots, \mathcal{S}_\beta$ during a visit in time frame t.

Lemma 3.3 *Let $\mathcal{M}$ be an $(n, m, \tau, c, s, \{h_j\}^{t=1,\ldots,\tau}_{j=1,\ldots,m})$ dynamic multi-threshold metering scheme. Let $A = \{1, \ldots, \alpha\}$ be a set of clients indices and let $B = \{1, \ldots, \beta\}$ be a set of servers indices. Then, for any time frame $t = 1, \ldots, \tau$, it holds that*

$$H(\mathbf{C}^t_{A,B}|\mathbf{C}_A) = 0.$$

Proof. We have that

$$\begin{aligned}
H(\mathbf{C}^t_{A,B}|\mathbf{C}_A) &\leq \sum_{i=1}^{\alpha}\sum_{j=1}^{\beta} H(\mathbf{C}^t_{i,j}|\mathbf{C}_A) \text{ (from (12) of Appendix)} \\
&\leq \sum_{i=1}^{\alpha}\sum_{j=1}^{\beta} H(\mathbf{C}^t_{i,j}|\mathbf{C}_i) \text{ (from (11) of Appendix)} \\
&= 0 \text{ (from Property 1 of Definition 2.1).}
\end{aligned}$$

□

The next lemma will be a useful tool to prove a lower bound on the size of the information distributed to servers from clients during a visit.

Lemma 3.4 *Let $\mathcal{M}$ be an $(n, m, \tau, c, s, \{h_j\}^{t=1,\ldots,\tau}_{j=1,\ldots,m})$ dynamic multi-threshold metering scheme. Let $\mathcal{S}_1 \ldots, \mathcal{S}_\beta$ be a coalition of $\beta \leq s$ corrupt servers and let $B = \{1, \ldots, \beta\}$. Let $\mathcal{C}_i$ be a client and for $j = 1, \ldots, \beta$ and $t = 1, \ldots, \tau$, let $X^t_{j,(h^t_j-1)}$ be a set of visits from $h^t_j - 1$ clients other than $\mathcal{C}_i$ to server $\mathcal{S}_j$ in time frame t. Then, for any $t = 1, \ldots, \tau$ and $i = 1, \ldots, n$, it holds that*

$$H(\mathbf{C}^t_{i,B}|\mathbf{X}^t_{1,(h^t_1-1)} \ldots \mathbf{X}^t_{\beta,(h^t_\beta-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B) \geq H(\mathbf{P}^t_B).$$

Proof. Let $\mathcal{C}_1, \ldots, \mathcal{C}_\alpha$ be a coalition of $\alpha \leq c$ corrupt servers other than $\mathcal{C}_i$. We have that

$$\begin{aligned}
&H(\mathbf{C}^t_{1,B} \ldots \mathbf{C}^t_{\alpha,B}|\mathbf{C}_1 \ldots \mathbf{C}_\alpha\mathbf{C}^t_{i,B}\mathbf{X}^t_{1,(h^t_1-\alpha-1)} \ldots \mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B) \\
&\quad\leq H(\mathbf{C}^t_{1,B} \ldots \mathbf{C}^t_{\alpha,B}|\mathbf{C}_1 \ldots \mathbf{C}_\alpha) \text{ (from (11) of Appendix)} \\
&\quad= 0 \text{ (from Lemma 3.3).}
\end{aligned}$$

Applying Lemma 3.2 with $\mathbf{A} = \mathbf{C}^t_{1,B} \ldots \mathbf{C}^t_{\alpha,B}$, $\mathbf{E} = \mathbf{C}_1 \ldots \mathbf{C}_\alpha\mathbf{C}^t_{i,B}\mathbf{X}^t_{1,(h^t_1-\alpha-1)} \ldots \mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B$, and $\mathbf{G} = \mathbf{P}^t_B$ we get

$$\begin{aligned}
&H(\mathbf{P}^t_B|\mathbf{C}_1 \ldots \mathbf{C}_\alpha\mathbf{C}^t_{i,B}\mathbf{X}^t_{1,(h^t_1-\alpha-1)} \ldots \mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B) \\
&\quad= H(\mathbf{P}^t_B|\mathbf{C}^t_{1,B} \ldots \mathbf{C}^t_{\alpha,B}\mathbf{C}_1 \ldots \mathbf{C}_\alpha\mathbf{C}^t_{i,B}\mathbf{X}^t_{1,(h^t_1-\alpha-1)} \ldots \mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B) \\
&\quad\leq H(\mathbf{P}^t_B|\mathbf{C}^t_{1,B} \ldots \mathbf{C}^t_{\alpha,B}\mathbf{C}^t_{i,B}\mathbf{X}^t_{1,(h^t_1-\alpha-1)} \ldots \mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B) \text{ (from (11) of Appendix)} \\
&\quad= H(\mathbf{P}^t_B|\mathbf{X}^t_{1,(h^t_1)} \ldots \mathbf{X}^t_{\beta,(h^t_\beta)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B) \\
&\quad\leq \sum_{j=1}^{\beta} H(\mathbf{P}^t_j|\mathbf{X}_{j,(h^t_j)}\mathbf{S}^t_j) \text{ (from (11) and (12)of Appendix)} \\
&\quad= 0 \text{ (from Property 2 of Definition 2.1).}
\end{aligned}$$

From Property 3 of Definition 2.1 we have that

$$H(\mathbf{P}^t_{B}|\mathbf{C}_1\dots\mathbf{C}_\alpha\mathbf{X}^t_{1,(h^t_1-\alpha-1)}\cdots\mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B) = H(\mathbf{P}^t_B).$$

Therefore, applying Lemma 3.2 with $\mathbf{E} = \mathbf{C}_1\dots\mathbf{C}_\alpha\mathbf{X}^t_{1,(h^t_1-\alpha-1)}\cdots\mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B$, $\mathbf{F} = \mathbf{C}^t_{i,B}$, and $\mathbf{G} = \mathbf{P}^t_B$, we get

$$\begin{aligned}
&H(\mathbf{C}^t_{i,B}|\mathbf{C}_1\dots\mathbf{C}_\alpha\mathbf{X}^t_{1,(h^t_1-\alpha-1)}\cdots\mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B)\\
&\qquad = H(\mathbf{P}^t_B) + H(\mathbf{C}^t_{i,B}|\mathbf{C}_1\dots\mathbf{C}_\alpha\mathbf{X}^t_{1,(h^t_1-\alpha-1)}\cdots\mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B\mathbf{P}^t_B)\\
&\qquad \geq H(\mathbf{P}^t_B) \text{ (from (4) of Appendix).} \qquad (1)
\end{aligned}$$

We have that

$$\begin{aligned}
H(\mathbf{C}^t_{1,B}\dots\mathbf{C}^t_{\alpha,B}|\mathbf{C}_1\dots\mathbf{C}_\alpha\mathbf{X}^t_{1,(h^t_1-\alpha-1)}\cdots\mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B\mathbf{P}^t_B) &\leq H(\mathbf{C}^t_{1,B}\dots\mathbf{C}^t_{\alpha,B}|\mathbf{C}_1\dots\mathbf{C}_\alpha)\\
&= \text{(from (11) of Appendix)}\\
&= 0 \text{ (from Lemma 3.3).}
\end{aligned}$$

Applying Lemma 3.1 with $\mathbf{A} = \mathbf{C}^t_{1,B}\dots\mathbf{C}^t_{\alpha,B}$, $\mathbf{E} = \mathbf{C}_1\dots\mathbf{C}_\alpha\mathbf{X}^t_{1,(h^t_1-\alpha-1)}\cdots\mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B$, and $\mathbf{G} = \mathbf{C}^t_{i,B}$ we get

$$\begin{aligned}
&H(\mathbf{C}^t_{i,B}|\mathbf{C}_1\dots\mathbf{C}_\alpha\mathbf{X}^t_{1,(h^t_1-\alpha-1)}\cdots\mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B)\\
&\quad = H(\mathbf{C}^t_{i,B}|\mathbf{C}^t_{1,B}\dots\mathbf{C}^t_{\alpha,B}\mathbf{C}_1\dots\mathbf{C}_\alpha\mathbf{X}^t_{1,(h^t_1-\alpha-1)}\cdots\mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B)\\
&\quad \leq H(\mathbf{C}^t_{i,B}|\mathbf{C}^t_{1,B}\dots\mathbf{C}^t_{\alpha,B}\mathbf{X}^t_{1,(h^t_1-\alpha-1)}\cdots\mathbf{X}^t_{\beta,(h^t_\beta-\alpha-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B) \text{ (from (11) of Appendix)}\\
&\quad = H(\mathbf{C}^t_{i,B}|\mathbf{X}^t_{1,(h^t_1-1)}\cdots\mathbf{X}^t_{\beta,(h^t_\beta-1)}\mathbf{S}^t_B\mathbf{V}^{[t-1]}_B) \qquad (2)
\end{aligned}$$

Therefore, the lemma follows from inequalities (2) and (1). $\square$

The next corollary immediately follows from Lemma 3.4. It provides a lower bound on the size of the information distributed to servers from clients during a visit in dynamic multi-threshold metering schemes. It states that in any dynamic multi-threshold metering scheme the size of the information $c^t_{i,j}$ given by client $\mathcal{C}_i$ to server $\mathcal{S}_j$ during a visit in time frame t is lower bounded by $\log|F|$, where F is the set in which the audit agency chooses the proofs for all the servers. This bound is tight, as in Section 4 we present a protocol for a $(n, m, \tau, c, s, \{h_j\}^{t=1,\dots,\tau}_{j=1,\dots,m})$ dynamic multi-threshold metering schemes in which the clients distribute *exactly* this information to servers.

Corollary 3.5 *Let $\mathcal{M}$ be an $(n, m, \tau, c, s, \{h_j\}^{t=1,\dots,\tau}_{j=1,\dots,m})$ dynamic multi-threshold metering scheme. For any $i = 1, \dots, n$, $j = 1, \dots, m$, and $t = 1, \dots, \tau$, it holds that*

$$H(\mathbf{C}^t_{i,j}) \geq H(\mathbf{P}^t_j).$$

If the proofs for the servers are uniformly chosen in a finite field F then Corollary 3.5 implies

$$H(\mathbf{C}^t_{i,j}) \geq \log|F|.$$

In order to prove a lower bound on the size of the information distributed to clients we need the next lemma.

Lemma 3.6 *Let $\mathcal{M}$ be an $(n, m, \tau, c, s, \{h_j\}_{j=1,\ldots,m}^{t=1,\ldots,\tau})$ dynamic multi-threshold metering scheme. Let $\mathcal{S}_1 \ldots, \mathcal{S}_\beta$ be a coalition of $\beta \leq s$ corrupt servers and let $B = \{1, \ldots, \beta\}$. Let $Z \subseteq \{1, \ldots, n\}$ be a set of clients indices. Then, it holds that*

$$H(\mathbf{C}_Z) \geq \sum_{t=1}^{\tau} H(\mathbf{C}^t_{Z,B} | \mathbf{V}_B^{[t-1]}).$$

Proof. We have that

$$\begin{aligned} H(\mathbf{C}^1_{Z,B} \ldots \mathbf{C}^\tau_{Z,B} | \mathbf{C}_Z) &\leq \sum_{t=1}^{\tau} H(\mathbf{C}^t_{Z,B} | \mathbf{C}_Z) \text{ (from (12) and (11) of Appendix)} \\ &= 0 \text{ (from Lemma 3.3)}. \end{aligned}$$

Therefore, applying Lemma 3.2 with $\mathbf{F} = \mathbf{C}^1_{Z,B} \ldots \mathbf{C}^\tau_{Z,B}$ and $\mathbf{D} = \mathbf{C}_Z$ we get

$$\begin{aligned} H(\mathbf{C}_Z) &= H(\mathbf{C}^1_{Z,B} \ldots \mathbf{C}^\tau_{Z,B}) + H(\mathbf{C}_Z | \mathbf{C}^1_{Z,B} \ldots \mathbf{C}^\tau_{Z,B}) \\ &\geq H(\mathbf{C}^1_{Z,B} \ldots \mathbf{C}^\tau_{Z,B}) \text{ (from (4) of Appendix)} \\ &= H(\mathbf{C}^1_{Z,B}) + \sum_{t=2}^{\tau} H(\mathbf{C}^t_{Z,B} | \mathbf{C}^1_{Z,B} \ldots \mathbf{C}^{t-1}_{Z,B}) \text{ (from (5) of Appendix)} \\ &\geq \sum_{t=1}^{\tau} H(\mathbf{C}^t_{Z,B} | \mathbf{V}_B^{[t-1]}). \end{aligned}$$

□

The next lemma provides a lower bound on the size of the information distributed to clients during the initialization phase in dynamic multi-threshold metering schemes. It states that in any dynamic multi-threshold metering scheme the size of the information c_i distributed to any client $\mathcal{C}_i$ by the audit agency is lower bounded by $s\tau \log |F|$, where F is the set in which the audit agency chooses the proofs for all the servers. This bound is tight, as in Section 4 we present a protocol for $(n, m, \tau, c, s, \{h_j\}_{j=1,\ldots,m}^{t=1,\ldots,\tau})$ dynamic multi-threshold metering schemes which distributes *exactly* this information to clients.

Lemma 3.7 *Let $\mathcal{M}$ be an $(n, m, \tau, c, s, \{h_j\}_{j=1,\ldots,m}^{t=1,\ldots,\tau})$ dynamic multi-threshold metering scheme. For any $i = 1, \ldots, n$ it holds that*

$$H(\mathbf{C}_i) \geq \sum_{t=1}^{\tau} H(\mathbf{P}^t_B).$$

Proof. Let $\mathcal{C}_i$ be a client and for $j = 1, \ldots, \beta$ and $t = 1, \ldots, \tau$, let $X^t_{j,(h^t_j-1)}$ be a set of visits from $h^t_j - 1$ clients other than $\mathcal{C}_i$ to server $\mathcal{S}_j$ in time frame t. We have that

$$\begin{aligned} H(\mathbf{C}_i) &\geq \sum_{t=1}^{\tau} H(\mathbf{C}^t_{i,B} | \mathbf{V}_B^{[t-1]}) \text{ (from Lemma 3.6)} \\ &\geq \sum_{t=1}^{\tau} H(\mathbf{C}^t_{i,B} | \mathbf{X}^t_{1,(h^t_1-1)} \ldots \mathbf{X}^t_{\beta,(h^t_\beta-1)} \mathbf{S}^t_B \mathbf{V}_B^{[t-1]}) \text{ (from (11) of Appendix)} \\ &\geq \sum_{t=1}^{\tau} H(\mathbf{P}^t_B) \text{ (from Lemma 3.4)}. \end{aligned}$$

□

Notice that in Definition 2.1 we did not say anything on the entropies of random variables $\mathbf{P}_j^t$ for $j \in \{1,\ldots,m\}$ and $t \in \{1,\ldots,\tau\}$. Our results apply to the general case of arbitrary entropies on proofs, but for clarity, in the following, we state our results for the simpler case that $H(\mathbf{P}_{j_1}^{t_1}) = H(\mathbf{P}_{j_2}^{t_2})$ for all $j_1, j_2 \in \{1,\ldots,m\}$ and $t_1, t_2 \in \{1,\ldots,\tau\}$. We denote this common entropies by $H(\mathbf{P})$.

If the proof sequences of the corrupt servers are statistically independent, then the next corollary holds.

Corollary 3.8 *Let $\mathcal{M}$ be an $(n, m, \tau, c, s, \{h_j\}_{j=1,\ldots,m}^{t=1,\ldots,\tau})$ dynamic multi-threshold metering scheme and let $\mathcal{S}_1,\ldots,\mathcal{S}_s$ be the s corrupt servers. If the sequences $\mathbf{P}_1^t,\ldots,\mathbf{P}_s^t$ of the proofs of the servers $\mathcal{S}_1,\ldots,\mathcal{S}_s$ are statistically independent, it holds that*

$$H(\mathbf{C}_i) \geq s\tau H(\mathbf{P}).$$

If the proofs for the servers are uniformly chosen in a finite field F then Corollary 3.8 implies

$$H(\mathbf{C}_i) \geq s\tau \log |F|.$$

4 A Dynamic Multi–Threshold Metering Protocol

In this section we present a protocol for dynamic multi–threshold metering schemes which achieves the bounds of Corollary 3.5 and Corollary 3.8. The protocol is a generalization of Naor and Pinkas metering scheme [7].

Initialization:
For $j = 1,\ldots,m$ and $t = 1,\ldots,\tau$, let h_j^t be the threshold associated to the server $\mathcal{S}_j$ in time frame t and let $h_M = \max_{j=1,\ldots,m} \max_{t=1,\ldots,\tau} h_j^t + 1$. The audit agency A chooses a random polynomial $Q(x,y)$ of degree $h_M - 1$ in x and $s\tau - 1$ in y over $GF(q)$, where q is a sufficiently large prime number. Afterwards, A sends to each client $\mathcal{C}_i$ the univariate polynomial $Q(i,y)$, which is of degree $s\tau - 1$.

Beginning of a Time Frame:
At the beginning of time frame t, for any server $\mathcal{S}_j$, the audit agency A evaluates the polynomial $Q(x, j \circ t)$ in $h_M - h_j^t$ points other than $1,\ldots,n$ and sends these values to $\mathcal{S}_j$. The argument $j \circ t$ denotes the concatenation of j and t, and we assume for simplicity that $j \circ t$ is in $GF(q)$ and that no distinct two pairs (j,t) and (j',t') are mapped to the same element.

Regular Operation:
When the client $\mathcal{C}_i$ visits the server $\mathcal{S}_j$ in time frame t, it sends to $\mathcal{S}_j$ the value $Q(i, j \circ t)$.

Proof Generation:
Assume that the server $\mathcal{S}_j$ has been visited by at least h_j^t different clients in time frame t. Then, knowing the $h_M - h_j^t$ points of $Q(x, j \circ t)$ given by the audit agency to $\mathcal{S}_j$ at the beginning of time frame t, it can perform a Lagrange interpolation and reconstruct the polynomial $Q(x, j \circ t)$. Then, it can compute the value $Q(0, j \circ t)$, which constitutes the proof that the server sends to the audit agency. The audit agency can easily verify this value.

4.1 Security of the Scheme

In this section we prove that the scheme presented in Section 4 satisfies Properties 1, 2, and 3 of Definition 2.1.

It is immediate to verify that the scheme satisfies Property 1 of Definition 2.1. Indeed, for any $i = 1, \ldots, n$, the information given by the audit agency to the client $\mathcal{C}_i$ consists of the univariate polynomial $Q(i, y)$ and for any $j = 1, \ldots, m$ and $t = 1, \ldots, \tau$, the information given to the server $\mathcal{S}_j$ by client $\mathcal{C}_i$ in time frame t is obtained by evaluating the univariate polynomial $Q(i, y)$ at $j \circ t$.

It is also easy to verify that the scheme satisfies Property 2 of Definition 2.1. Assume that a server $\mathcal{S}_j$ has been visited by h_j^t clients in time frame t and that it has received $h_M - h_j^t$ points of $Q(x, j \circ t)$ from the audit agency at the beginning of time frame t Therefore, the server $\mathcal{S}_j$ knows h_M points of the polynomial $Q(x, j \circ t)$ and can perform a Lagrange interpolation on it. Afterwards, it can compute its proof $Q(0, j \circ t)$ by evaluating the polynomial $Q(x, j \circ t)$ in the point 0.

Finally, we prove that the scheme satisfies Property 3 of Definition 2.1. Consider a coalition of $\beta \leq s \leq m$ corrupt servers and $\alpha \leq c < \min_{j=1,\ldots,m} \min_{t=1,\ldots,\tau} h_j^t$ corrupt clients, which decide to cooperate at a certain time frame t. Assume that during time frame t each server $\mathcal{S}_j$ in the coalition has received less than $h_j^t - \alpha$ regular visits from clients. We will prove that the servers in the coalition have no information on their proofs for time frame t. We consider the worst possible case in which the coalition of corrupt clients and servers has the maximum size c and s, respectively. Moreover we assume that the corrupt servers and clients decide to cooperate at time frame τ and that corrupt servers have collected the maximum possible information during the previous time frames $1, \ldots, \tau - 1$. In other words, we assume that each corrupt client $\mathcal{C}_i$ donates its polynomial $Q(i, y)$ to all servers in the coalition, and that any corrupt server $\mathcal{S}_j$ in the coalition knows the polynomial $Q(x, j \circ t)$ for $t = 1, \ldots, \tau - 1$. This means that in time frames $t = 1, \ldots, \tau - 1$ each server $\mathcal{S}_j$ in the coalition has computed its proof as function of its client visits and of the information provided by the audit agency.

In order to compute its proof $Q(0, j \circ \tau)$ for time frame τ, any server $\mathcal{S}_j$ should be able to interpolate either the polynomial $Q(x, j \circ \tau)$ or the bivariate polynomial $Q(x, y)$. Therefore, we consider the two following cases:

Case 1. The server $\mathcal{S}_j$ attempts to interpolate the polynomial $Q(x, j \circ \tau)$ in order to compute its proof.

Notice that for any $j, k \in \{1, \ldots, s\}$, with $j \neq k$, the information held by the server $\mathcal{S}_k$ is of no help in computing the polynomial $Q(x, j \circ \tau)$. Assume $g_j = h_j^\tau - c - 1$ be the number of client visits received by server $\mathcal{S}_j$ during time frame τ. Each corrupt client $\mathcal{C}_i$ donates to $\mathcal{S}_j$ the polynomial $Q(i, y)$ from which $\mathcal{S}_j$ can compute the value $Q(i, j \circ \tau)$. Since there are c corrupt clients, $\mathcal{S}_j$ can compute c values of $Q(x, j \circ \tau)$ in addition to those provided by the g_j visits performed by non corrupt clients. Since the server $\mathcal{S}_j$ has also received $h_M - h_j^\tau$ points of $Q(x, j \circ \tau)$ by the audit agency at the beginning of time frame τ, the overall number of points of $Q(x, j \circ \tau)$ known to $\mathcal{S}_j$ is $g_j + c + h_M - h_j^\tau = h_M - 1$. Therefore, the server obtains a linear system of $h_M - 1$ equations in h_M unknowns. For any choice of a value in $GF(q)$, there is a polynomial $R(x, j \circ \tau)$ which is consistent with the information held by the server. Since there are q such polynomials, the probability of the server in guessing its proof for time frame τ is at most $1/q$.

Case 2. The coalition of servers attempt to interpolate the polynomial $Q(x, y)$ in order to compute the proofs.

The information that a corrupt client $\mathcal{C}_i$ donates to a corrupt server is equivalent to the $s\tau$ coefficients of its polynomial $Q(i, y)$. For $j = 1, \ldots, s$, the information collected by each corrupt server $\mathcal{S}_j$ at the beginning of time frame τ is constituted by the information provided by the audit agency at the beginning of any time frame $t = 1, \ldots, \tau$, which consists in $h_M - h_j^t$ coefficients of $Q(x, j \circ t)$, plus the information provided by clients during each time frame $t = 1, \ldots, \tau - 1$, which consists in h_j^t coefficients of $Q(x, j \circ t)$. Hence, at the beginning of time frame τ each corrupt server helds $(\tau - 1)h_M$ coefficients of $Q(x, y)$ and $h_M - h_j^\tau$ coefficients of $Q(x, j \circ \tau)$. Suppose that in time frame τ each server $\mathcal{S}_j$, $j \in \{1, \ldots, s\}$, receives $g_j \leq h_j^\tau - \alpha - 1$ regular visits form clients. Then, the overall information on $Q(x, y)$ held by the coalition of corrupt servers and clients at the end of time frame τ consists of

$$cs\tau + s(\tau - 1)h_M + \sum_{j=1}^{s}(h_M - h_j^\tau) + \sum_{j=1}^{s} g_j - cs(\tau - 1) \tag{3}$$

points. The first term of (3) is the information donated by the c corrupt clients, the second term is the information collected by the s corrupt servers during time frames $1, \ldots, \tau - 1$, the third term is the information provided by the audit agency at the beginning of time frame τ, the fourth term is the information provided by client visits at time frame τ, and the last term is the information which has been counted twice. Since $g_j \leq h_j^\tau - \alpha - 1$ for $j = 1, \ldots, s$ then it is easy to see that expression (3) is less than or equal to $h_M s\tau - s$. Therefore, the servers obtain a system of $h_M s\tau - s$ equations in $h_M s\tau$ unknowns. For any choice of s values in $GF(q)$, there is a polynomial $R(x, y)$ which is consistent with the information held by the servers in the coalition Since there are q^s such polynomials, then the corrupt servers $\mathcal{S}_1, \ldots, \mathcal{S}_s$ have probability at most $1/q^s$ of guessing their proofs for time frame τ.

4.2 Efficiency of the Scheme

In this section we analyze the efficiency of the scheme presented in Section 4. It is easy to see that the scheme meets the bounds of Corollary 3.8 and Corollary 3.5. Indeed, during the initialization phase each client $\mathcal{C}_i$ receives by the audit agency the polynomial $Q(i, y)$, which is of degree $s\tau - 1$. Therefore, the size of the information distributed to any client is $s\tau \log q$ and the bound of Corollary 3.8 is tight. During a regular operation in a time frame t each client $\mathcal{C}_i$ gives to the visited server $\mathcal{S}_j$ the value $Q(i, j \circ t)$. Therefore, the size of the information distributed to any visited server is $\log q$ and the bound of Corollary 3.5 is tight. Then, this protocol is optimal both with respect to the size of information distributed to clients and with respect to the size of information given to servers by clients. this is important otherwise the task of receiving and sending information would burden the clients, that are not interested in the metering process.

5 Conclusions and Open Problems

In this paper we have introduced dynamic multi-threshold metering schemes. In these schemes the servers need to communicate with the audit agency at the beginning of any time frame.

In order to devise efficient dynamic multi-threshold metering schemes, we have decided to minimize the information distribution involving clients. This is important because clients are not interested in the metering process, i.e., they do not receive any money from the audit agency. Therefore, we have only considered the problem of bounding the size of the information distributed to clients and the size of the information distributed from clients to servers during the visits.

An interesting problem would be to provide lower bounds on the size of the information distributed to servers at the beginning of any time frame and to devise dynamic multi-threshold metering schemes in which this information is as small as possible.

Acknowledgements

This work was done while the third author was visiting the Department of Combinatorics and Optimization at the University of Waterloo, Ontario, Canada. She wants to thank the Department for its hospitality. The research of the fourth author is supported by the Natural Sciences and Research Council of Canada under grants NSERC-IRC 216431-96 and NSERC-RGPIN 203114-98.

References

[1] V. Anupam, A. Mayer, K. Nissim, B. Pinkas, M. K. Reiter, *On the Security of Pay-Per-Click and Other Web Advertising Schemes*, 8th International World Wide Web Conference.

[2] C. Blundo, A. De Bonis and B. Masucci, *Metering Schemes with Pricing*, submitted for publication, 2000.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley & Sons, 1991.

[4] A. De Bonis and B. Masucci, *An Information Theoretical Approach to Metering Schemes*, accepted at 2000 International Symposium on Information Theory, June 2000, Italy.

[5] M. Franklin and D. Malkhi, *Auditable Metering with Lightweight Security*, in "Financial Cryptography '97", Lecture Notes in Computer Science, Vol. **1318**, pp. 151–160, 1997.

[6] B. Masucci and D. R. Stinson *Efficient Metering Schemes with Pricing*, submitted for publication, 2000. Technical Report CORR 2000-06, Centre for Applied Cryptographic Research, University of Waterloo.

[7] M. Naor and B. Pinkas, *Secure and Efficient Metering*, in "Advances in Cryptology - EUROCRYPT '98", Lecture Notes in Computer Science, Vol. **140**, pp. 576–590, 1998.

[8] M. Naor and B. Pinkas, *Secure Accounting and Auditing on the Web*, Computer Networks and ISDN Systems, Vol. **40**, Issues 1-7, pp. 541–550, 1998.

[9] T. P. Novak and D. L. Hoffman, *New Metrics for New Media: Toward the Development of Web Measurement Standards.* Available on line at `http://www2000.ogsm.vanderbilt.edu/novak/ web.standards/webstand.html`

[10] A. Shamir, *How to Share a Secret,* Commun. of the ACM, Vol. **22**, pp. 612–613, 1979.

Appendix

Information Theory Background

In this Appendix we review the basic concepts of Information Theory used in our definitions and proofs. For a complete treatment of the subject the reader is advised to consult [3].

Given a probability distribution $\{Pr_{\mathbf{X}}(x)\}_{x\in X}$ on a set X, we define the *entropy* [1] of $\mathbf{X}$, $H(\mathbf{X})$, as

$$H(\mathbf{X}) = -\sum_{x\in X} Pr_{\mathbf{X}}(x)\log Pr_{\mathbf{X}}(x).$$

The entropy $H(\mathbf{X})$ is a measure of the average uncertainty one has about which element of the set X has been chosen when the choices of the elements from X are made according to the probability distribution $\{Pr_{\mathbf{X}}(x)\}_{x\in X}$. It is well known that $H(\mathbf{X})$ is a good approximation to the average number of bits needed to faithfully represent the elements of X. The entropy satisfies the following property

$$0 \leq H(\mathbf{X}) \leq |X|,$$

where $H(\mathbf{X}) = 0$ if and only if there exists $x_0 \in X$ such that $Pr_{\mathbf{X}}(x_0) = 1$; whereas, $H(\mathbf{X}) = \log|X|$ if and only if $Pr_{\mathbf{X}}(x) = 1/|X|$ for all $x \in X$.

Given two sets X and Y and a joint probability distribution on their cartesian product, the *conditional entropy* $H(\mathbf{X}|\mathbf{Y})$, is defined as

$$H(\mathbf{X}|\mathbf{Y}) = -\sum_{y\in Y}\sum_{x\in X} Pr_{\mathbf{Y}}(y)Pr(x|y)\log Pr(x|y).$$

From the definition of conditional entropy it is easy to see that

$$H(\mathbf{X}|\mathbf{Y}) \geq 0. \tag{4}$$

We have that $H(\mathbf{X}|\mathbf{Y}) = 0$ when the value taken from Y completely determines the value chosen from X; whereas, $H(\mathbf{X}|\mathbf{Y}) = H(\mathbf{X})$ means that choices from X and Y are independent, that is, the probability that the value x has been chosen from X given that from Y we have chosen y is the same as the *a priori* probability of choosing x from X. Therefore, knowing the values chosen from Y does not enable a Bayesian opponent to modify an *a priori* guess regarding which element has been chosen from X.

Given $n+1$ sets $X_1,\ldots,X_n,Y$ and a joint probability distribution on their cartesian product, the entropy of $\mathbf{X}_1\ldots\mathbf{X}_n$ given $\mathbf{Y}$ can be expressed as

$$H(\mathbf{X}_1\ldots\mathbf{X}_n|\mathbf{Y}) = H(\mathbf{X}_1|\mathbf{Y}) + \sum_{i=2}^{n} H(\mathbf{X}_i|\mathbf{X}_1\ldots\mathbf{X}_{i-1}\mathbf{Y}). \tag{5}$$

The *mutual information* $I(\mathbf{X};\mathbf{Y})$ between $\mathbf{X}$ and $\mathbf{Y}$ is defined by

$$I(\mathbf{X};\mathbf{Y}) = H(\mathbf{X}) - H(\mathbf{X}|\mathbf{Y}) \tag{6}$$

and enjoys the following properties:

$$I(\mathbf{X};\mathbf{Y}) = I(\mathbf{Y};\mathbf{X}) \tag{7}$$

[1] All log's in this paper denote basis 2 logarithms.

and $I(\mathbf{X};\mathbf{Y}) \geq 0$, from which one gets

$$H(\mathbf{X}) \geq H(\mathbf{X}|\mathbf{Y}). \tag{8}$$

Given three sets X, Y, Z and a joint probability distribution on their cartesian product, the *conditional mutual information* $I(\mathbf{X};\mathbf{Y}|\mathbf{Z})$ between $\mathbf{X}$ and $\mathbf{Y}$ given $\mathbf{Z}$ is

$$I(\mathbf{X};\mathbf{Y}|\mathbf{Z}) \;=\; H(\mathbf{X}|\mathbf{Z}) - H(\mathbf{X}|\mathbf{Z}\mathbf{Y}) \tag{9}$$

and enjoys the following properties:

$$I(\mathbf{X};\mathbf{Y}|\mathbf{Z}) = I(\mathbf{Y};\mathbf{X}|\mathbf{Z}) \tag{10}$$

and $I(\mathbf{X};\mathbf{Y}|\mathbf{Z}) \geq 0$. Since the conditional mutual information is always non negative we get

$$H(\mathbf{X}|\mathbf{Z}) \geq H(\mathbf{X}|\mathbf{Z}\mathbf{Y}). \tag{11}$$

From (5) and (11) one easily gets that for any sets $Y, X_1, \ldots, X_n$ and a joint probability distribution on their cartesian product it holds that

$$\sum_{i=1}^{n} H(\mathbf{X}_i|\mathbf{Y}) \geq H(\mathbf{X}_1\mathbf{X}_2\ldots\mathbf{X}_n|\mathbf{Y}). \tag{12}$$

n	number of clients
m	number of servers
τ	number of time frames
c	number of corrupt clients
s	number of corrupt servers
h_j^t	threshold for server $\mathcal{S}_j$ in time frame t
$\mathbf{C}_i$	information distributed to client $\mathcal{C}_i$
$\mathbf{C}_{i,j}^t$	visit from client $\mathcal{C}_i$ to server $\mathcal{S}_j$ in time frame t
$B = \{1, \ldots, \beta\}$	indices of corrupt servers, $\beta \leq s$
$\mathbf{C}_{i,B}^t$	visits from client $\mathcal{C}_i$ to servers $\mathcal{S}_1, \ldots, \mathcal{S}_\beta$ in time frame t
$\mathbf{X}_{j,(d_j)}^t$	visits from d_j clients to server $\mathcal{S}_j$ in time frame t
$\mathbf{S}_j^t$	information distributed to server $\mathcal{S}_j$ at the beginning of time frame t
$\mathbf{S}_B^t$	information distributed to servers $\mathcal{S}_1, \ldots, \mathcal{S}_\beta$ at the beginning of time frame t
$\mathbf{P}_j^t$	proof for server $\mathcal{S}_j$ in time frame t
$\mathbf{P}_B^t$	proofs for servers $\mathcal{S}_1, \ldots, \mathcal{S}_\beta$ in time frame t
$\mathbf{V}_j^{[t]}$	information collected by server $\mathcal{S}_j$ in time frames $1, \ldots, t$
$\mathbf{V}_B^{[t]}$	information collected by servers $\mathcal{S}_1, \ldots, \mathcal{S}_\beta$ in time frames $1, \ldots, t$

Figure 1. Parameters and variables frequently used in the paper.