SUT Journal of Mathematics
Vol. 46, No. 2 (2010), 243–254

On the average number of ciphertexts for the optimal key-management tree for a frequency distribution of communication

Nozomu Ochiumi

(Received August 6, 2010; Revised November 24, 2010)

Abstract. The average number of ciphertexts per one session $\bar{c}_T$ is an indicator of aptitude of a key-management tree T for a given frequency distribution of communication. In this paper, we estimate the weighted mean of $\bar{c}_T$ over all T's with respect to appropriate weights as a criterion of aptitude of a key-management tree. We also give the concrete value of that mean in the case of equal weights.

AMS 2010 *Mathematics Subject Classification.* 68R10, 05C05, 94A60.

Key words and phrases. Broadcast encryption schemes, 1-to-k encryption, tree-based key-management system, frequency distribution of communication, average cost, entropy, generating function method.

§1. Introduction

1.1. 1-to-k cryptography and tree-based key-management system

The 1-to-k cryptography is a recent problem arising in a scheme of broadcast communications, which is motivated largely by pay-TV applications, multicast communications, secure distribution of copyright-protected material (e.g. music) and audio streaming. A 1-to-k encryption involves 1 transmitter and n users. Each user is given some keys, and the transmitter is given all the keys by the key manager. Let U be the whole set of n users and J ($\subseteq U$) be a set of k authorized receivers. The transmitter sends ciphertexts to all users, but any receiver in J should be able to decrypt the received ciphertext by using the receiver's key while users in the revoked set $U \setminus J$ should not be able to do so. The algorithm for the 1-to-k encryption-decryption consists of the following steps.

Initialization (by the key manager)

1. Set a family of subsets of users $\mathcal{F}_n = \{S_1, \ldots, S_m\}$ $(\subseteq 2^U)$ satisfying $\bigcup_{i=1}^m S_i = U$ and $\binom{U}{1} \subseteq \mathcal{F}_n$.
2. For each $i = 1, \ldots, m$, create a key K_i and transmit it to all the users $\in S_i$ by a secret communication.

Encryption/Transmission (by the transmitter)

1. Choose a set of receivers J $(\subseteq U)$ to whom the message M is addressed.
2. Take a family of disjoint subsets $\{S_{i_1}, \ldots, S_{i_t}\}$ $(\subseteq \mathcal{F}_n)$ satisfying $\bigcup_{l=1}^t S_{i_l} = J$ in such a way as to minimize t.
3. Encrypt the message M by using each key $K_{i_1}, \ldots, K_{i_t}$ and transmit the ciphertexts $C_{i_1}, \ldots, C_{i_t}$ to all the users.

Reception/Decryption (by each receiver $u \in J$)

1. Find j satisfying $u \in S_{i_j}$.
2. Decrypt the ciphertext C_{i_j} to get the message M by using the key K_{i_j} which u owns.

We consider two quantities as main criteria for performance evaluations; the number of ciphertexts for a set of receivers, the number of keys owned by a user. A problem to be considered is how to set a family of subsets of users $\mathcal{F}_n$. If $\mathcal{F}_n$ consists of all the non-empty subsets of U (i.e., $\mathcal{F}_n = \bigcup_{i=1}^n \binom{U}{i}$), the transmitter always sends 1 ciphertext and each user must have 2^{n-1} keys, so that this method is not suitable for users with limited memory capacities. If $\mathcal{F}_n$ consists of all 1-subsets of users (i.e., $\mathcal{F}_n = \binom{U}{1}$), each user has only 1 key and the transmitter always sends k different ciphertexts, so that this method is not suitable for channels with limited bandwidth. Thus the number of keys which one user should own and the number of ciphertexts which should be sent to users are in a trade-off relationship. Many methods of keeping that balance have been proposed. One of them is the tree-based key-management system suggested by Wong et al. [7] and Naor et al. [3]

In a tree-based key-management system, each user corresponds to a leaf of a rooted tree T in one-to-one manner and S_i corresponds to a node of the tree. Hereafter T is assumed to be binary. The most popular method is to assign each node v the set S_v of users corresponding to leaves under v and set $\mathcal{F}_n = \{S_v \mid v \text{ is a node of } T\}$. Then each user owns the keys assigned to the nodes on the path from the root to the leaf corresponding to that user. The number of keys owned by a user and the number of ciphertexts for a set of

receivers are uniquely determined once we have chosen the tree structure T and the placement of users among its leaves. Henceforth we regard the number of necessary ciphertexts as cost. Let $c_T(J)$ be the cost for a set of receivers J in a key-management tree T (define $c_T(\emptyset) = 0$).

Note that J corresponds to a subset of leaves in a key-management tree T with user-set U. If we delete all the paths (and all the edges incident to them) that connect the root and the leaves in the revoked set $U \setminus J$, then there remains a forest consisting of $c_T(J)$ number of subtrees of T. The $c_T(J)$ is called in [4] the *covering number* for J in T.

1.2. Average cost under a frequency distribution of communication

In this paper we take into consideration a frequency distribution of communication in the tree-based key-management system. Let $p(J)$ be the probability that a message is sent to $J \in 2^U \setminus \{\emptyset\}$. In the case that users are not too many, the transmitter may estimate the frequency distribution of communication $\big(p(J) \bigm| J \in 2^U \setminus \{\emptyset\}\big)$ by tallying up the number of communications to receiver set J. For a given frequency distribution of communication $(p(J))$, the average cost is given by

$$\bar{c}_T = \sum_{J \in 2^U \setminus \{\emptyset\}} p(J) c_T(J).$$

A tree T minimizing this value is called optimal, and denoted by T^*.

Let $\mathcal{T}_U$ denote the set of all the key-management trees for the set of users U. A key-management tree can be regarded as a graphic representation of a binary total partition of U. Put $b_n = |\mathcal{T}_U|$, then it is known that $b_1 = 1$,

$$b_t = \frac{1}{2} \sum_{s=1}^{t-1} \binom{t}{s} b_s b_{t-s} = (2t-3)!!, \ t \geqslant 2, \tag{1.1}$$

where $b_0 = 0$ (see [6]). $t!!$ means the double factorial, that is, $(-1)!! = 0!! = 1$ and $t!! = t(t-2)!!$ for $t \geqslant 1$. Therefore the total number of key-management trees is $O(t^t)$, so that it is not easy to find an optimal tree by an exhaustive search. Funayama et al. [1] and Imamura et al. [2] proposed algorithms for generating key-management trees suitable for a given frequency distribution of communication. Both algorithms, however, do not always generate optimal trees for some distributions, and need exhaustive searches of all trees to ascertain the optimality. For this reason, Ochiumi et al. [4, 5] introduced an upper bound on the average cost of the optimal tree T^* as a criterion of aptitude of a key-management tree for a given frequency distribution of communication.

Theorem A ([4, 5]). *Let* $a_{n,k} = \sum_{T\in\mathcal{T}_U} c_T(J)$ *for* $|U| = n$ *and* $J \in \binom{U}{k}$. *Then*

$$\bar{c}_{T^*} \leqslant \frac{1}{(2n-3)!!}\sum_{k=1}^{n} a_{n,k}P\{|J| = k\}$$

holds for $n \geqslant 2$. *And* $a_{n,k}$ *satisfies the following formula:*

$$a_{n,k} = (2(n-k)-1)!!\left(\frac{(2n-2)!!}{(2(n-k)-2)!!} - \frac{(2n-3)!!}{(2(n-k)-3)!!}\right), 1 \leqslant k \leqslant n-1.$$

It was also shown in [4] that the polygonal line though the points $\left(\frac{k}{n}, \frac{a_{n,k}}{(2n-1)b_n}\right)$ $(0 \leqslant k \leqslant n)$ approaches the curve $f(\rho) = \sqrt{1-\rho}(1-\sqrt{1-\rho})$ when $n \to \infty$ with fixing $\frac{k}{n}$ to ρ.

The purpose of this paper is to estimate a lower bound on the average cost for the optimal tree. For a given frequency distribution of communication, good key-management trees are expected to be balanced if the distribution is flat, and imbalanced if it is biased. It seems natural to think that the cost for the optimal tree is close to the weighted mean of $\bar{c}_T$ with respect to Ts' weights w_T ($w_T \geqslant 0$ and $\sum_{T\in\mathcal{T}_U} w_T = 1$) when we put high weights on trees considered to be suitable for a key-management tree.

This paper is organized as follows. In section 2, we propose a lower bound on $\sum_{T\in\mathcal{T}_U} w_T\bar{c}_T$ in terms of the Shannon entropy $H(p(J))$ of the frequency distribution of communication. We also give a neat formula for the lower bound when the weights of the trees are uniform. In section 3, we state about some remarks for the lower bound.

§2. Results

The weighted mean of the average costs over all the trees satisfies the following inequality.

Proposition 1.

$$\sum_{T\in\mathcal{T}_U} w_T\bar{c}_T \geqslant H(p(J)) - \log\sum_{J\in 2^U\setminus\{\emptyset\}}\sum_{T\in\mathcal{T}_U} w_T 2^{-c_T(J)}. \tag{2.1}$$

Proof. For the given frequency distribution of communication $(p(J))$, we have

$$H(p(J)) = -\sum_{J\in 2^U\setminus\{\emptyset\}} p(J)\log p(J) \leqslant -\sum_{J\in 2^U\setminus\{\emptyset\}} p(J)\log\frac{2^{-c_T(J)}}{\sum_{J'\in 2^U\setminus\{\emptyset\}} 2^{-c_T(J')}}$$

by the log-sum inequality. Hence we obtain

$$\bar{c}_T = \sum_{J\in 2^U\setminus\{\emptyset\}} p(J)c_T(J) \geqslant H(p(J)) - \log \sum_{J\in 2^U\setminus\{\emptyset\}} 2^{-c_T(J)}.$$

And we have

$$\begin{aligned}\sum_{T\in\mathcal{T}_U} w_T\bar{c}_T &\geqslant H(p(J)) - \sum_{T\in\mathcal{T}_U} w_T\left(\log \sum_{J\in 2^U\setminus\{\emptyset\}} 2^{-c_T(J)}\right)\\ &\geqslant H(p(J)) - \log \sum_{J\in 2^U\setminus\{\emptyset\}}\sum_{T\in\mathcal{T}_U} w_T 2^{-c_T(J)},\end{aligned}$$

where the second inequality follows from the Jensen's inequality. □

Let w_T depend only on T's form and be independent of the way in which users correspond to leaves in T. Then $\sum_{T\in\mathcal{T}_U} w_T 2^{-c_T(J)}$ depends only on $|J|$ and with respect to (2.1) we obtain

$$\sum_{J\in 2^U\setminus\{\emptyset\}}\sum_{T\in\mathcal{T}_U} w_T 2^{-c_T(J)} = \sum_{k=1}^{n}\sum_{J\in\binom{U}{k}}\sum_{T\in\mathcal{T}_U} w_T 2^{-c_T(J)} = \sum_{k=1}^{n}\binom{n}{k}\frac{\alpha_{n,k}}{b_n}, \tag{2.2}$$

where $\alpha_{n,k} = b_n\sum_{T\in\mathcal{T}_U} w_T 2^{-c_T(J)}$ for $|U| = n$, $J\in\binom{U}{k}$ $(n\geqslant 1,\ 0\leqslant k\leqslant n)$ and b_n is given by (1.1). Particularly, in the case that the weights of the trees are uniform, that is, $w_T = b_n^{-1}$ for all $T\in\mathcal{T}_U$,

$$\alpha_{n,k} = \sum_{T\in\mathcal{T}_U} 2^{-c_T(J)},\ |U| = n,\ J\in\binom{U}{k}\ (n\geqslant 1,\ 0\leqslant k\leqslant n). \tag{2.3}$$

In Theorem 2, we obtain an explicit formula for the general term $\alpha_{n,k}$ in that case.

Theorem 2. *Let $\alpha_{n,k}$ be defined as in* (2.3). *Then for $n\geqslant 1$ and $0\leqslant k\leqslant n$,*

$$\alpha_{n,k} = \frac{1}{2^k}(2(n-k)-1)!!\frac{(2n-2)!}{(2n-k-1)!}. \tag{2.4}$$

We need two lemmas to give a proof of the theorem.

Lemma 3. *For all $l, m\in\mathbb{N}$,*

$$\sum_{i=0}^{l-1} C_{l-i-1}\binom{m+2i+1}{i} = \binom{m+2l}{l-1}, \tag{2.5}$$

where C_t $(t\geqslant 0)$ is the t-th Catalan number (i.e., $C_t = \frac{1}{t+1}\binom{2t}{t}$).

Proof. We shall prove that (2.5) holds for all $m \in \mathbb{N}$ by induction on l. (2.5) clearly holds for $l = 1$. For $l \geqslant 1$, we have

$$\begin{aligned}
&\sum_{i=0}^{l} C_{l-i}\binom{m+2i+1}{i} \\
&\qquad = \binom{m+2l+1}{l} + \sum_{i=0}^{l-1} C_{l-i}\binom{m+2i+1}{i} \\
&\qquad = \binom{m+2l+1}{l} + \sum_{i=0}^{l-1}\left(\sum_{j=i+1}^{l} C_{l-j}C_{j-i-1}\right)\binom{m+2i+1}{i} \\
&\qquad = \binom{m+2l+1}{l} + \sum_{j=1}^{l} C_{l-j}\sum_{i=0}^{j-1} C_{j-i-1}\binom{m+2i+1}{i} \\
&\qquad = \binom{m+2l+1}{l} + \sum_{j=1}^{l} C_{l-j}\binom{m+2j}{j-1} \\
&\qquad = \binom{m+2l+1}{l} + \binom{m+2l+1}{l-1} \\
&\qquad = \binom{m+2l+2}{l},
\end{aligned}$$

where the second equality holds by the recursion of the Catalan numbers

$$C_{t+1} = \sum_{s=0}^{t} C_s C_{t-s}$$

and the fourth and the fifth by the induction hypothesis, that is,

$$\sum_{i=0}^{j-1} C_{j-i-1}\binom{m+2i+1}{i} = \binom{m+2j}{j-1} \text{ for } 1 \leqslant j \leqslant l$$

and

$$\sum_{j=1}^{l} C_{l-j}\binom{m+2j}{j-1} = \sum_{j=0}^{l-1} C_{l-j-1}\binom{m+1+2j+1}{j} = \binom{m+1+2l}{l-1}.$$

Hence (2.5) holds for all $l, m \in \mathbb{N}$. □

Lemma 4. *For all $m \in \mathbb{N}$,*

$$\left(\frac{1}{2}\left(1+\sqrt{1-4y}\right)\right)^{-m} = \sum_{l\geqslant 0} \frac{m}{m+2l}\binom{m+2l}{l} y^l, \ |y| < \frac{1}{4}. \tag{2.6}$$

Proof. For $m = 1$, we have

$$\begin{aligned} 2\left(1+\sqrt{1-4y}\right)^{-1} &= \frac{1}{2y}\left(1-\sqrt{1-4y}\right) \\ &= \frac{1}{2y}\left(\sum_{l\geqslant 1}\frac{(2l-3)!!}{l!}2^l y^l\right) \\ &= \sum_{l\geqslant 1}\frac{1}{2l-1}\binom{2l-1}{l-1}y^{l-1}, \end{aligned}$$

hence (2.6) holds. Assume that (2.6) holds for some $m \geqslant 1$. Differentiate the both sides of (2.6) with respect to y, then we have

$$\begin{aligned} \left(\frac{1}{2}\left(1+\sqrt{1-4y}\right)\right)^{-(m+1)} &= \sqrt{1-4y}\sum_{l\geqslant 1}\frac{l}{m+2l}\binom{m+2l}{l}y^{l-1} \\ &= \sum_{i\geqslant 0}\binom{\frac{1}{2}}{i}(-4y)^i\sum_{l\geqslant 0}\binom{m+2l+1}{l}y^l \\ &= \sum_{l\geqslant 0}\sum_{i=0}^{l}\binom{\frac{1}{2}}{l-i}(-4)^{l-i}\binom{m+2i+1}{i}y^l, \end{aligned}$$

where $\binom{\lambda}{t} = \dfrac{\lambda(\lambda-1)\cdots(\lambda-t+1)}{t!}$ for $t \in \mathbb{N}$ and $\binom{\lambda}{0} = 1$. Here we see that

$$\sum_{i=0}^{l}\binom{\frac{1}{2}}{l-i}(-4)^{l-i}\binom{m+2i+1}{i} = \frac{m+1}{m+2l+1}\binom{m+2l+1}{l},$$

because it is clear for $l = 0$, and

$$\begin{aligned} &\sum_{i=0}^{l}\binom{\frac{1}{2}}{l-i}(-4)^{l-i}\binom{m+2i+1}{i} \\ &\qquad = -2\sum_{i=0}^{l-1}\frac{1}{l-i}\binom{2(l-i-1)}{l-i-1}\binom{m+2i+1}{i} + \binom{m+2l+1}{l} \\ &\qquad = -2\binom{m+2l}{l-1} + \binom{m+2l+1}{l} \qquad \text{(by Lemma 3)} \\ &\qquad = \frac{m+1}{m+2l+1}\binom{m+2l+1}{l} \end{aligned}$$

for $l \geqslant 1$. Hence (2.6) holds for $m+1$. Therefore (2.6) holds for all $m \in \mathbb{N}$. □

Proof of Theorem 2. It is clear that $\alpha_{n,0} = (2n-3)!! = b_n$ and $\alpha_{n,n} = \frac{1}{2}(2n-3)!! = \frac{1}{2}b_n$ since $c_T(\emptyset) = 0$ and $c_T(U) = 1$ for all $T \in \mathcal{T}_U$. For $1 \leqslant k \leqslant n-1$, let T_1, T_2 be the two subtrees of the root of $T \in \mathcal{T}_U$. And define V and its complement V^c as the sets of leaves of T_1 and T_2 respectively. Then $c_T(J) = c_{T_1}(J \cap V) + c_{T_2}(J \cap V^c)$ holds for $J \in \binom{U}{k}$ $(1 \leqslant k \leqslant n-1)$, and we have

$$\begin{aligned}
\alpha_{n,k} &= \sum_{T \in \mathcal{T}_U} 2^{-c_T(J)} \\
&= \frac{1}{2} \sum_{l=1}^{n-1} \sum_{V \in \binom{U}{l}} \sum_{T_1 \in \mathcal{T}_V} \sum_{T_2 \in \mathcal{T}_{V^c}} 2^{-\{c_{T_1}(J \cap V) + c_{T_2}(J \cap V^c)\}} \\
&= \frac{1}{2} \sum_{l=1}^{n-1} \sum_{V \in \binom{U}{l}} \sum_{T_1 \in \mathcal{T}_V} 2^{-c_{T_1}(J \cap V)} \sum_{T_2 \in \mathcal{T}_{V^c}} 2^{-c_{T_2}(J \cap V^c)} \\
&= \frac{1}{2} \sum_{l=1}^{n-1} \sum_{V \in \binom{U}{l}} \alpha_{l,|J \cap V|} \alpha_{n-l,|J \cap V^c|}.
\end{aligned}$$

Noting that $|J \cap V| = i$ iff $|J \cap V^c| = k - i$ and $\left|\left\{V \in \binom{U}{l} \mid |J \cap V| = i\right\}\right| = \left|\left\{V \in \binom{U}{l} \mid |J \cap V^c| = k - i\right\}\right| = \binom{k}{i}\binom{n-k}{l-i}$ for $0 \leqslant i \leqslant k$, we obtain

$$\begin{aligned}
\alpha_{n,k} &= \frac{1}{2} \sum_{l=1}^{n-1} \sum_{i=0}^{k} \binom{k}{i} \binom{n-k}{l-i} \alpha_{l,i} \alpha_{n-l,k-i} \\
&= \frac{1}{2} \sum_{i=0}^{k} \sum_{l=i}^{n-k+i} \binom{k}{i} \binom{n-k}{l-i} \alpha_{l,i} \alpha_{n-l,k-i},\ 1 \leqslant k \leqslant n-1,
\end{aligned}$$

where $\alpha_{0,0} = 0$. Putting $A_{n,k} = \dfrac{\alpha_{n,k}}{(n-k)!k!}$ for $0 \leqslant k \leqslant n$ and $f_k(x) = \sum_{n \geqslant k} A_{n,k} x^{n-k}$ for $k \geqslant 0$, we have

$$A_{n,k} = \frac{1}{2} \sum_{i=0}^{k} \sum_{l=i}^{n-k+i} A_{l,i} A_{n-l,k-i}$$

for $1 \leqslant k \leqslant n-1$. Hence we have

$$\begin{aligned} f_k(x) - A_{k,k} &= \sum_{n \geqslant k+1} A_{n,k} x^{n-k} \\ &= \sum_{n \geqslant k+1} \left(\frac{1}{2} \sum_{i=0}^{k} \sum_{l=i}^{n-k+i} A_{l,i} A_{n-l,k-i} \right) x^{n-k} \\ &= \frac{1}{2} \sum_{i=0}^{k} \left(\sum_{n \geqslant k} \sum_{l=i}^{n-k+i} \left(A_{l,i} x^{l-i} \right) \left(A_{n-l,k-i} x^{n-l-k+i} \right) - A_{i,i} A_{k-i,k-i} \right) \\ &= \frac{1}{2} \sum_{i=0}^{k} \left(f_i(x) f_{k-i}(x) - A_{i,i} A_{k-i,k-i} \right), \text{ for } k \geqslant 1. \end{aligned}$$

Putting $f(x,y) = \sum_{k \geqslant 0} f_k(x) y^k = \sum_{k \geqslant 0} \sum_{n \geqslant k} A_{n,k} x^{n-k} y^k$, we have

$$\begin{aligned} f(x,y) - f_0(x) - \sum_{k \geqslant 0} A_{k,k} y^k &= \frac{1}{2} \sum_{k \geqslant 1} \sum_{i=0}^{k} \left(f_i(x) f_{k-i}(x) - A_{i,i} A_{k-i,k-i} \right) y^k \\ &= \frac{1}{2} \left(f(x,y)^2 - f_0(x)^2 - \left(\sum_{k \geqslant 0} A_{k,k} y^k \right)^2 \right). \end{aligned}$$

Since $A_{n,0} = \frac{b_n}{n!}$, $A_{k,k} = \frac{1}{2} \frac{b_k}{k!}$ and $\sum_{t \geqslant 0} \frac{b_t}{t!} z^t = 1 - \sqrt{1-2z}$, we have

$$\begin{aligned} (f(x,y) - 1)^2 &= \left(\sum_{k \geqslant 0} A_{k,k} y^k - 1 \right)^2 + (f_0(x) - 1)^2 - 1 \\ &= \frac{1}{4} \left(1 + \sqrt{1-2y} \right)^2 - 2x, \end{aligned}$$

and

$$\begin{aligned}
&f(x,y)\\
&= 1-\frac{1}{2}\sqrt{\left(1+\sqrt{1-2y}\right)^2-8x}\\
&= 1-\frac{1}{2}\left(1+\sqrt{1-2y}\right)\sum_{t\geqslant 0}\binom{\frac{1}{2}}{t}(-8x)^t\left(1+\sqrt{1-2y}\right)^{-2t}\\
&= 1-\frac{1}{2}\left(1+\sqrt{1-2y}\right)+\sum_{t\geqslant 1}\frac{(2t-3)!!}{t!}x^t\left(\frac{1}{2}\left(1+\sqrt{1-2y}\right)\right)^{-(2t-1)}\\
&= \frac{1}{2}\left(1-\sqrt{1-2y}\right)+\sum_{t\geqslant 1}\frac{(2t-3)!!}{t!}x^t\sum_{k\geqslant 0}\frac{2t-1}{2t+2k-1}\binom{2t+2k-1}{k}\left(\frac{y}{2}\right)^k\\
&= \frac{1}{2}\sum_{n\geqslant 0}\frac{b_n}{n!}y^n+\sum_{k\geqslant 0}\sum_{t\geqslant 1}\frac{(2t-1)!!(2t+2k-2)!}{2^k t!k!(2t+k-1)!}x^t y^k,
\end{aligned}$$

where the fourth equality follows from Lemma 4. Substituting t with $n-k$, we have

$$A_{n,k}=\frac{1}{2^k}\frac{(2(n-k)-1)!!(2n-2)!}{(n-k)!k!(2n-k-1)!},\ 0\leqslant k\leqslant n-1.$$

From the above, the proof is complete. □

§3. Concluding remarks

Using (2.2) and (2.4) in (2.1), when the weights of the trees are uniform, we get the following inequality:

$$\frac{1}{|\mathcal{T}_U|}\sum_{T\in\mathcal{T}_U}\bar{c}_T\geqslant H(p(J))-\log\sum_{k=1}^{n}\binom{n}{k}\frac{1}{2^k}\frac{(2(n-k)-1)!!(2n-2)!}{(2n-k-1)!(2n-3)!!}. \tag{3.1}$$

The second term on the right-hand side of (2.1) is determined by the weights of the trees, which may be taken suitably for the given frequency distribution of communication. An effective lower bound of $\sum_{T\in\mathcal{T}_U}w_T\bar{c}_T$ could be derived by choosing appropriate the weights of the trees. The right-hand side of (3.1) could be negative and then (3.1) would make no sense, which might be caused by setting the weights of the trees uniform regardless of the given frequency distribution of communication. Calculating the values of $\log\sum_{k=1}^{n}\binom{n}{k}\frac{\alpha_{n,k}}{b_n}$ for $n=1,\ldots,1000$ and plotting the points of $\left(n,\log\sum_{k=1}^{n}\binom{n}{k}\frac{\alpha_{n,k}}{b_n}\right)$, we get the following figure.

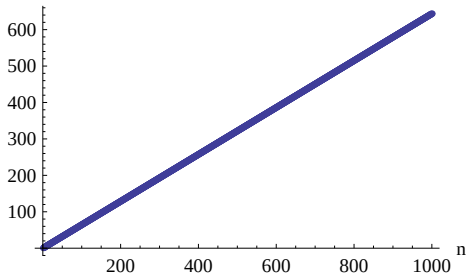

Figure 1: $\log \sum_{k=1}^{n} \binom{n}{k} \frac{\alpha_{n,k}}{b_n}$ for $n = 1, \ldots, 1000$

From this calculation, we expect

$$\frac{1}{n} \log \sum_{k=1}^{n} \binom{n}{k} \frac{\alpha_{n,k}}{b_n} \fallingdotseq 0.6438$$

even when n is large. In the special case that each user is contained in the receivers set J independently with probability p, the entropy of the frequency distribution of communication is $nH(p, 1-p)$ ($H(p, 1-p)$: the binary entropy). When $0.17 \leqslant p \leqslant 0.83$, $H(p, 1-p) \geqslant 0.6438$ and the right-hand side of (3.1) is positive.

We call $c_T(J)$ the covering number for J in tree T ([4]). A simple lower bound for the expected covering number provided by the completely random choices of T (with n leaves) and J would be about $0.356n$, since we have $H(p(J)) = \log(2^n - 1)$.

Acknowledgements

The author would like to thank the Associate Editor and the referee for their useful comments. And the author would like to express his sincere thanks to Professor Yasuichi Horibe and Professor Masahiro Yanagida for their useful suggestions and comments.

References

[1] T. Funayama, S. Imamura, and E. Okamoto, *Efficient key distribution system using communication probability*, IPSJ SIG Technical Report, 2006-CSEC-33 **2006** (2006), no. 43, pp. 1–6.

[2] S. Imamura, Y. Oyama, T. Okamoto, and E. Okamoto, *Construction of binary key management trees based on frequency distribution of communication*, Proceedings

of the 2007 Symposium on Cryptography and Information Security (SCIS2007), 2007, p. 98.

[3] D. Naor, M. Naor, and J. Lotspiech, *Revocation and tracing schemes for stateless receivers*, Advances in cryptology—CRYPTO 2001 (Santa Barbara, CA), Lecture Notes in Comput. Sci. **2139** (2001), pp. 41–62.

[4] N. Ochiumi, F. Kanazawa, M. Yanagida, and Y. Horibe, *On the average number of nodes covering a given number of leaves in an unordered binary tree*, to appear in Journal of Combinatorial Mathematics and Combinatorial Computing.

[5] N. Ochiumi, F. Kanazawa, M. Yanagida, E. Okamoto, and Y. Horibe, *Criteria of aptitude of a key-management tree for a frequency distribution of communication*, Proceedings of the 2008 International Symposium on Information Theory and Its Applications (ISITA 2008), 2008, pp. 201–203.

[6] R. P. Stanley, *Enumerative combinatorics. Vol. 2*, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999.

[7] C. K. Wong, M. G. Gouda, and S. S. Lam, *Secure group communications using key graphs*, IEEE/ACM Trans. Networking **8** (2000), no. 1, pp. 16–30.

Nozomu Ochiumi
Department of Mathematical Information Science, Tokyo University of Science
1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan
E-mail: ochiumi@hotmail.co.jp