Distribution of the Number of Encryptions in Revocation Schemes for Stateless Receivers †

Christopher Eagle[1] and Zhicheng Gao[2] and Mohamed Omar[3] and Daniel Panario[2] and Bruce Richmond[1]

[1] *Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
email: {*cjeagle@engmail.uwaterloo.ca, lbrichmond@math.uwaterloo.ca*}
[2] *School of Mathematics and Statistics, Carleton University, Ottawa, Ontario K1S 5B6, Canada*
email: {*zgao@math.carleton.ca, daniel@math.carleton.ca*}
[3] *Department of Mathematics, University of California Davis, Davis, California 95616, United States of America*
email: {*momar@math.ucdavis.edu*}

We study the number of encryptions necessary to revoke a set of users in the complete subtree scheme (CST) and the subset-difference scheme (SD). These are well-known tree based broadcast encryption schemes. Park and Blake in: Journal of Discrete Algorithms, vol. 4, 2006, pp. 215–238, give the mean number of encryptions for these schemes. We continue their analysis and show that the limiting distribution of the number of encryptions for these schemes is normal. This implies that the mean numbers of Park and Blake are good estimates for the number of necessary encryptions used by these schemes.

Keywords: Key distribution schemes, subset-difference scheme (SD), complete subtree scheme (CST), broadcast encryption, Hwang's quasi-power theorem, Heuberger's two dimensions quasi-power theorem.

1 Introduction

We consider the problem of a center broadcasting an encrypted message to a group of users such that some subset is revoked, and hence should not be able to obtain the content of the broadcasted message even if they collaborate. Various encryption schemes have been proposed to solve this practical problem which arises in pay-TV, satellite communications, real-time information update and media content protection. In one class of proposed schemes, the center distributes a unique combination of keys to each user, and users decrypt the message individually. If keys cannot be updated once distributed, the receivers are called *stateless*. The keys are distributed so that no revoked (or excluded) user has a decryption key and every privileged (non-revoked) user has at least one decryption key. If the subset of privileged users is arbitrary and dynamically changing, one is faced with the problem of minimizing the number of encryptions necessary to ensure system security.

† The authors have been partially funded by NSERC of Canada.

To achieve the above goals, several key distribution schemes use a balanced binary tree data structure implementation. Two examples that we consider are the *subset-difference scheme* (SD), introduced by Naor, Naor and Lotspiech (6), and the *complete subtree scheme* (CST), introduced independently by Wallner, Harder, Agee (8) and Wong, Gouda and Lam (9).

In the CST scheme each user is represented as a unique leaf node in a balanced binary tree. Every node, including the leaves, is assigned a unique decryption key, and each user holds the keys which are on the path from its leaf node to its root node. A user can decrypt a transmission if he or she holds at least one key with which the transmission was encrypted. When a user is revoked, all keys in the path from its leaf to the root become unavailable. As users dynamically change, the number and actual decryption keys available also change. Given a set of users, we are interested in the minimum number of encryption keys that the center needs to broadcast in order to guarantee that every privileged user in the system is able to decrypt the transmission (and every revoked user is not able to decrypt it). This scheme was introduced in (8; 9) and can be adapted to the stateless case (that is, when keys are assigned by the broadcasting center and they cannot be updated during the operation of the system). A more detailed explanation of the CST scheme, and how the algorithm handles the revocation of users while still guaranteeing that every privileged user has at least one decryption key, can be found in Section 3.1 of (6) and in Section 2.1 of (7).

A modification of the complete subtree scheme is the subset-difference scheme (SD). In this scheme, the users are again considered as leaves of a complete binary tree, but the key assignment scheme is altered. First, for any vertex i of the binary tree, and any descendant j of the subtree rooted at i, we define the set $S_{i,j}$ to be the set of leaves (i.e. users) of the complete binary tree that are descendants of i but not descendants of j. Then, $S_{i,j}$ is naturally called a *subset difference*, and to each subset difference set, the broadcaster assigns a key $K_{i,j}$. Hence, any leaf u will be assigned the collection of keys $\{K_{i,j}\}$ where i, j runs through all pairs of vertices in the tree such that i is an ancestor of u but j is not an ancestor of u.

The broadcaster associates an m bit label for each vertex i in the complete binary tree. The broadcaster then recursively assigns keys $K_{i,j}$ for every descendant j of i as follows: first, a psuedo random generator takes as input the m bit label x given to i and outputs a $3m$ bit label x_{new}. The first m bits of x_{new} is then a temporary m bit label of the left child of j, the last m bits of x_{new} is a temporary m bit label of the right child of j, and $K_{i,j}$ is the middle m bits of x_{new}. This process is then continued inductively with the temporary labels of the children of j.

The appealing property of the subset difference scheme is that given a specific key $K_{i,j}$ for the subset difference $S_{i,j}$, one can easily determine the key for the subset difference $S_{i,j'}$ if j' is a descendant of j, and vice-versa. Revoking users is also relatively easy. Given a set of revoked users $\mathfrak{R}$, the strategy to revoke users is to simply send out keys only to subset differences $S_{i,j}$ where $S_{i,j} \cap \mathfrak{R} = \emptyset$. An efficient implementation of this is given in Naor, Naor and Lotspiech (6).

A practical application of the work of Naor, Naor and Lotspiech (6) is found in the Blu-ray technology. This technology, which is the successor of the DVD technology, relies on the Advanced Access Content System (AACS) for its security features (1). AACS was developed by several leading technology and media companies and includes the use of Naor, Naor and Lotspiech's SD scheme for the encryption of the media content (see (2), Section 3.2.1). In the case of the subset-difference scheme by AACS each device capable of decoding the Blu-ray device is treated as a user in the system, so it is possible to block compromised playback devices from viewing future releases by revoking the corresponding user in the SD scheme.

Yet another revocation scheme is the *layered subset-difference scheme* (LSD), which is a modification of the subset-difference scheme. Instead of assigning keys to all subset differences $S_{i,j}$, the LSD scheme restricts to subset differences $S_{i,j}$ where i, j are relatively close to each other. In particular, the vertices of the complete binary tree which users are leaves of are partitioned into classes by their distance from the root (i.e. their level). There are $\sqrt{\log_2 n}$ partition classes each with $\sqrt{\log_2 n}$ levels, and levels $k \cdot \sqrt{\log_2 n}$ where $k = 0, 1, \ldots, \sqrt{\log_2 n}$ are called *special* levels. Subset differences $S_{i,j}$ are used in the LSD scheme if and only if both i and j lie in the same partition class or i is at a special level. As shown by Halevy and Shamir (3) (Lemma 2), every subset difference $S_{i,j}$ in the SD scheme is the disjoint union of at most two valid subset differences in the LSD scheme. The revocation process for this scheme is the same as that of the SD scheme. Though revocation may seem more difficult for a broadcaster using the LSD scheme (because the number of sets needed to cover a set of revoked users doubles at worst), an advantage to using it is a significant reduction in the number of keys any user needs to hold (see Halevy and Shamir (3), Lemma 3).

In (7), Park and Blake give generating functions that entail the exact mean number of encryptions for the key distribution schemes CST and SD. They also considered the LSD scheme. We shall need the results of Park and Blake, and indeed, our analysis can be regarded as a continuation of their work. We show that the number of encryptions in all these schemes is asymptotically normally distributed. Our results imply that the average number of encryptions provided by Park and Blake are indeed very good estimates for the number of encryptions in these methods.

The structure of this paper is as follows. In Section 2, we review the results of Park and Blake (7). Our proofs require a two dimensional extension of Hwang's (5) quasi-power theorem due to Heuberger (4). We recall this result in Section 3. In Section 4, we give the main results of this paper: the limiting distributions of the number of encryptions for the CST and the SD schemes are normal. The number of encryptions for the LSD scheme is also normal but we do not include the proof here. We also give joint distributions for the number of encryptions and number of privileged users for the above schemes. Conclusions are given in Section 5.

2 Mean Number of Encryptions

We now start from the analysis and notation of Park and Blake (7). They suppose that there are $N = 2^n$ users in the system. It seems possible to generalize their generating functions to the case where $n = \lfloor \log_2 N \rfloor$, but for simplicity we only consider the case $N = 2^n$ users.

We observe that when we have $j \leq 2^n$ users to be served (privileged users), they could require any number of encryptions $i \leq j$. We denote by (i, j)-privileged users a set of j privileged users that require i encryptions. In the complete binary tree representation, a given set of privileged users can be partitioned into users in the left subtree and users in the right subtree. The number of (i, j)-privileged users in a system of 2^n users can be expressed as the number of (i', j')-privileged users in the left subtree in a system of 2^{n-1} users, and $(i - i', j - j')$-privileged users in the right subtree in a system of 2^{n-1} users. Let $a_{ij}^{(n)}$ denote the number of subsets of j privileged users which require exactly i encryptions. We have the generating function for the numbers $a_{ij}^{(n)}$

$$\sum_{j=0}^{2^n} \sum_{i=0}^{j} a_{ij}^{(n)} x^i y^j.$$

If there are $j^{'}$ users in the left subtree and $j - j^{'}$ users in the right subtree we have

$$a_{ij}^{(n)} = \sum_{j'=0}^{j} \sum_{i'=0}^{i} a_{i'j'}^{(n-1)} a_{i-i'\,j-j'}^{(n-1)}.$$

Using this recurrence, Park and Blake (7) give a recurrence for the generating function of the numbers $a_{ij}^{(n)}$ in the CST scheme.

Theorem 2.1 (Park-Blake) *The generating function for the CST scheme is*

$$\begin{aligned} T_0(x,y) &= 1 + xy, \\ T_n(x,y) &= T_{n-1}^2(x,y) + (1-x)xy^{2^n} \quad \textit{for} \;\; n \geq 1. \end{aligned}$$

In this generating function, y represents privileged users while x marks the number of encryptions. For the initial condition $n = 0$, the number of users is $N = 1$. It is clear that the set of no users require no encryptions, and one user requires one encryption, hence the expression $1 + xy$.

For $n \geq 1$, using the above partition of (i, j)-privileged users in a system of 2^n users in terms of the number of $(i^{'}, j^{'})$-privileged users in the left subtree in a system of 2^{n-1} users, and $(i-i^{'}, j-j^{'})$-privileged users in the right subtree in a system of 2^{n-1} users, we obtain the term $T_{n-1}^2(x, y)$. An adjusting term has to be added when the number of privileged users is $j = 2^n$ (all users are in the system) since in that case only one encryption is required, more precisely, the root key. Thus, the correct expression is xy^{2^n} and we must add the correcting term $(1 - x)xy^{2^n}$.

The generating function for the SD scheme can be obtained in a similar but more complicated derivation (see Section 3.2 in (7)).

Theorem 2.2 (Park-Blake) *The generating function for the SD scheme is*

$$\begin{aligned} S_0(x,y) &= 1 + xy, \\ S_n(x,y) &= S_{n-1}^2(x,y) + D_{n-1}(x,y) \qquad \textit{for} \;\; n \geq 1, \end{aligned}$$

where

$$\begin{aligned} D_0(x,y) &= (1-x)xy^2, \\ D_{n-1}(x,y) &= (1-x)x\left[y^{2^n} + 2^n y^{2^n} \sum_{i=0}^{n-2} 2^{-i} y^{-2^i}\right] \qquad \textit{for} \;\; 2 \leq n \leq 3, \end{aligned}$$

and, for $n \geq 4$, we have that $D_{n-1}(x, y)$ equals

$$(1-x)xy^{2^n}\left[1 + 2^n \sum_{i=0}^{1} 2^{-i} y^{-2^i} + 2^{n-1} \sum_{i=1}^{n-3} 2^{-i} y^{-2^{i+1}} \left(S_i(x,y) - xy^{2^i}\right)^2\right].$$

Park and Blake use the above generating functions to give exact expressions for the mean number of encryptions over all privileged sets for the considered schemes. Since we have N users there are 2^N

possible privileged sets of users. They assume that each of these 2^N possible privileged sets have the same probability. Then, the mean number of encryption is defined by

$$m(n) = \frac{\sum_j \sum_i i a_{ij}^{(n)}}{2^N} = \frac{1}{2^N} \frac{\partial G_n(x,y)}{\partial x}(1,1), \tag{1}$$

where $G_n(x,y)$ can be either $T_n(x,y)$ or $S_n(x,y)$, as defined in Theorems 2.1 and 2.2, respectively. They prove the following exact mean number estimates.

Theorem 2.3 (Park-Blake) *The mean number of encryptions over all privileged sets for the CST scheme is given by*

$$m_{\texttt{CST}}(n) = \frac{N}{2} - \left(\sum_{k=0}^{n-1} 2^{k-N2^{-k}} \right), \ n \geq 1,$$

with $m_{\texttt{CST}}(0) = 0.5$.

Theorem 2.4 (Park-Blake) *The mean number of encryptions over all privileged sets for the SD scheme is given by*

$$m_{\texttt{SD}}(n) = \frac{595N}{2048} - 13 \left(\sum_{i=0}^{n-4} 2^{i-N2^{-i}} \right) - \left(\sum_{i=0}^{n-4} N 2^{-N2^{-i}} \sum_{k=1}^{n-3-i} 2^{2^k - k} \right), \ n \geq 4,$$

with $m_{\texttt{SD}}(0) = 0.5$, $m_{\texttt{SD}}(1) = 0.75$, $m_{\texttt{SD}}(2) = 1.1875$ *and* $m_{\texttt{SD}}(3) = 2.324$.

We take the Park-Blake analysis a bit further by providing limiting distributions for the number of encryptions for these schemes. They also considered the LSD scheme providing a complicated generating function and its mean number of encryptions; see (7), Theorems 4 and 7. Our results also apply to this scheme, though the proof is not presented here.

Park and Blake derive asymptotic estimates for the means which they find to be accurate numerical estimates in comparison with the approximations given in (3; 6). Our results prove that their mean estimates are accurate estimates for the actual number of encryptions required in these schemes. Our results also provide precise variance and other moments for the number of encryptions.

3 Background

In this section we present a two dimensional version of a result of Hwang (5) due to Heuberger (4). Hwang's results, the so-called *quasi-power theorem*, give a central limit theorem and convergence rate for a sequence of random variables with moment generating function obeying a quasi-power form (that is, the moment generating function is asymptotically exponential). This theorem is useful in combinatorics since many combinatorial structures have asymptotic moment generating functions of this form.

Since in our problem we have two sequences of random variables (the number of encryptions and privileged users in a random privileged set), we require a bivariate version of Hwang's quasi-power theorem to deal with the joint distribution. We use the extension to two dimensions of Hwang's results due to Heuberger (4).

We use the following notation: $||(s,t)|| = \max\{|s|,|t|\}$; for a given function $u(s,t)$, we define

$$\mu_1 = \left.\frac{\partial u}{\partial s}\right|_{(0,0)}, \qquad \mu_2 = \left.\frac{\partial u}{\partial t}\right|_{(0,0)},$$

and

$$\sigma_1^2 = \left.\frac{\partial^2 u(s,t)}{\partial s^2}\right|_{(0,0)}, \quad \sigma_2^2 = \left.\frac{\partial^2 u(s,t)}{\partial t^2}\right|_{(0,0)}, \quad \sigma_{12} = \left.\frac{\partial^2 u(s,t)}{\partial s \partial t}\right|_{(0,0)};$$

finally, we denote by Σ the matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{bmatrix}.$$

We now state Heuberger (4) result that will be used to prove our theorems.

Theorem 3.1 (Heuberger) *Let $\{X_n, Y_n\}_{n\geq 1}$ be a sequence of two dimensional integral random vectors. Suppose that the moment generating function satisfies the asymptotic expression*

$$\begin{aligned} M_n(s,t) &= \sum_{m_1\geq 0, m_2 \geq 0} \mathbb{P}(X_n = m_1, Y_n = m_2) e^{m_1 s + m_2 t} \\ &= e^{u(s,t)\phi(n) + v(s,t)} \left(1 + O(1/\alpha_n)\right), \end{aligned}$$

where the O-term is uniform for $||(s,t)|| \leq \tau$, $(s,t) \in \mathbb{C}^2$, $\tau > 0$, and

(1) $u(s,t)$ and $v(s,t)$ are analytic for $||(s,t)|| \leq \tau$ and independent of n; the matrix Σ is nonsingular; and

(2) $\lim_{n\to\infty} \phi(n) = \infty$, and $\lim_{n\to\infty} \alpha_n = \infty$.

Then, the distribution of (X_n, Y_n) is asymptotically normal, i.e.,

$$\mathbb{P}\left(\frac{X_n - \mu_1 \phi(n)}{\sqrt{\phi(n)}} \leq x, \frac{Y_n - \mu_2\phi(n)}{\sqrt{\phi(n)}} \leq y\right) = \Phi_\Sigma(x,y) + O\left(\frac{1}{\sqrt{\phi(n)}} + \frac{1}{\alpha_n}\right),$$

where Φ_Σ denotes the two dimensional normal distribution with mean $(0,0)$ and covariance matrix Σ

$$\Phi_\Sigma(x_1, x_2) = \frac{1}{2\pi\sqrt{\det(\Sigma)}} \iint_{y_1 \leq x_1, y_2 \leq x_2} e^{-\frac{1}{2}(y_1,y_2)\Sigma^{-1}(y_1,y_2)^t} dy_1 dy_2.$$

4 Limiting distributions

Let X_n and Y_n, respectively, be random variables representing the number of encryptions and privileged users in a random privileged set. In this section we show that $\{X_n, Y_n\}_{n\geq 1}$ is asymptotically normal. We then, as a corollary, obtain that the marginal distributions of the number of encryptions and number of privileged users are also normally distributed.

More precisely we prove the following result.

Theorem 4.1 *With the above notation and for the CST and SD schemes, we have*

$$\mathbb{P}\left(\frac{X_n - 2^n\mu_1}{2^{n/2}} \le x, \frac{Y_n - 2^n\mu_2}{2^{n/2}} \le y\right) = \Phi_\Sigma(x,y)\left(1 + O\left(2^{-n/2}\right)\right),$$

where μ_1, μ_2 and the covariance matrix Σ are independent of n and can be computed efficiently.

We now prove in detail Theorem 4.1 for the CST scheme.

Lemma 4.2 *For all $n \ge 0$, $|x-1| \le 1/10$, and $|y-1| \le 1/10$, we have*

$$|T_n(x,y)| \ge (4/3)(4/3)^{2^n}.$$

PROOF. We use Theorem 2.1 and induction on n. It is clear that

$$\begin{aligned} |T_0(x,y)| &= |1+xy| = |2 + (x-1)(y-1) + (x-1) + (y-1)| \\ &\ge 2 - 1/100 - 1/5 > (4/3)^2. \end{aligned}$$

For the induction step, we have

$$\begin{aligned} |T_n(x,y)| &\ge |T_{n-1}^2(x,y)| - |(1-x)xy^{2^n}| \\ &\ge (4/3)^2(4/3)^{2^n} - (11/100)(11/10)^{2^n} \\ &= (4/3)(4/3)^{2^n} + (4/9)(4/3)^{2^n} - (11/100)(11/10)^{2^n} \\ &\ge (4/3)(4/3)^{2^n}. \end{aligned}$$

■

Lemma 4.3 *For all $n \ge 0$, $|x-1| \le 1/10$, $|y-1| \le 1/10$, $x = e^s$ and $y = e^t$, we have,*

$$T_n(e^s, e^t) = \exp\left(2^n u(s,t) + O\left((33/40)^{2^n}\right)\right), \tag{2}$$

where

$$u(s,t) = \ln(1+xy) + \sum_{j\ge 0} 2^{-j-1}\ln\left(1 + (1-x)xy^{2^{j+1}}T_j^{-2}(x,y)\right) \tag{3}$$

is an analytic function in a neighbor of $(s,t) = (0,0)$.

PROOF. In the following we assume $x = e^s$, $y = e^t$, and s and t are in a neighborhood of 0 such that $|x-1| < 1/10$ and $|y-1| \le 1/10$.

We derive an asymptotic estimate for $G_n(s,t) = \ln T_n(e^s, e^t)$. From Theorem 2.1, we obtain

$$\begin{aligned} G_n(s,t) &= 2G_{n-1}(s,t) + \ln\left(1 + (1-x)xy^{2^n}T_{n-1}^{-2}(x,y)\right) \\ &= 2^n\ln(1+xy) + \sum_{j=0}^{n-1} 2^{n-1-j}\ln\left(1 + (1-x)xy^{2^{j+1}}T_j^{-2}(x,y)\right) \end{aligned}$$

Since

$$(1-x)xy^{2^n}T_{n-1}^{-2}(x,y) = O\left((33/40)^{2^n}\right),$$

we have

$$\ln\left(1+(1-x)xy^{2^n}T_{n-1}^{-2}(x,y)\right)=O\left((33/40)^{2^n}\right).$$

Hence, the series

$$\sum_{j\geq 0}2^{-j-1}\ln\left(1+(1-x)xy^{2^{j+1}}T_j^{-2}(x,y)\right)$$

converges uniformly in a neighbor of $(x,y)=(1,1)$. This implies that

$$u(s,t)=\ln(1+xy)+\sum_{j\geq 0}2^{-j-1}\ln\left(1+(1-x)xy^{2^{j+1}}T_j^{-2}(x,y)\right)$$

is an analytic function in a neighbor of $(s,t)=(0,0)$. Thus we have

$$\begin{aligned} G_n(s,t) &= 2^n u(s,t)-\sum_{j\geq n}2^{n-j-1}\ln\left(1+(1-x)xy^{2^{j+1}}T_j^{-2}(x,y)\right)\\ &= 2^n u(s,t)+O\left((33/40)^{2^n}\right), \end{aligned}$$

and so we have derived Equation (2). ∎

PROOF OF THEOREM 4.1. For convenience, we define

$$A_j=\left.\frac{\partial T_j(x,y)}{\partial x}\right|_{(1,1)},\quad B_j=\left.\frac{\partial T_j(x,y)}{\partial y}\right|_{(1,1)},$$

and let us write T_j for $T_j(x,y)$. Then we have, from Equation (3),

$$\begin{aligned} \frac{\partial u}{\partial s} &= \frac{\partial x}{\partial s}\frac{\partial u}{\partial x}=x\left(\frac{y}{1+xy}\right.\\ &\quad\left.+\sum_{j\geq 0}2^{-j-1}y^{2^{j+1}}\frac{(1-2x)T_j^{-2}-2(1-x)xT_j^{-3}\frac{\partial T_j}{\partial x}}{1+(1-x)xy^{2^{j+1}}T_j^{-2}}\right)\\ \frac{\partial u}{\partial t} &= \frac{\partial y}{\partial t}\frac{\partial u}{\partial y}=y\left(\frac{x}{1+xy}\right.\\ &\quad\left.+\sum_{j\geq 0}2^{-j-1}(1-x)x\frac{2^{j+1}y^{2^{j+1}-1}T_j^{-2}-2y^{2^{j+1}}T_j^{-3}\frac{\partial T_j}{\partial y}}{1+(1-x)xy^{2^{j+1}}T_j^{-2}}\right), \end{aligned}$$

and consequently

$$\begin{aligned} \mu_1 &= \frac{1}{2}\left(1-\sum_{j\geq 0}2^{-j}4^{-2^j}\right),\\ \mu_2 &= \frac{1}{2}, \end{aligned}$$

$$\begin{aligned}
\sigma_1^2 &= \mu_1 - \frac{1}{4} + \sum_{j\geq 0} 2^{-j-1}\left(4\cdot 8^{-2^j}A_j - 2\cdot 4^{-2^j} - 16^{-2^j}\right),\\
\sigma_2^2 &= \mu_2 - \frac{1}{4} = \frac{1}{4},\\
\sigma_{1,2} &= \frac{1}{4} - \sum_{j\geq 0}\left(4^{-2^j} - 2^{-j}\,8^{-2^j}B_j\right).
\end{aligned}$$

The first few values of A_j and B_j can be computed using the following recursions

$$\begin{aligned}
A_0 &= B_0 = 1,\\
A_j &= 2^{1+2^{j-1}}A_{j-1} - 1,\\
B_j &= 2^{1+2^{j-1}}B_{j-1} - 1.
\end{aligned}$$

We note that all the series appearing above converge very rapidly as the general terms decrease doubly exponentially. By taking the first 4 terms (j from 0 to 3), we obtain the following values (correct to 9 decimal places)

$$\mu_1 = 0.358885765,\ \sigma_1^2 = 0.094122395,\ \sigma_{1,2} = 0.091789245.$$

Finally the proof of Theorem 4.1 for the CST scheme follows from Equation (2), Theorem 3.1, and the fact that the matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{bmatrix}$$

is nonsingular. ■

The following is an immediate corollary of Theorem 4.1 by taking the marginal distribution for X_n. It can also be obtained directly using Hwang's (5) quasi-power theorem.

Corollary 4.4 *Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables for the number of encryptions in the CST scheme. Then, X_n is asymptotically normal with mean $2^n\mu_1$ and variance $2^n\sigma_1^2$. More precisely, we have*

$$\mathbb{P}\left(\frac{X_n - 2^n\mu_1}{2^{n/2}\sigma_1} \leq x\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-s^2/2}ds + O\left(2^{-n/2}\right).$$

Clearly, this corollary provides all relevant statistical information about X_n. If we want, in addition to the exact mean provided by Park and Blake in Theorem 2.3, the *exact* variance of X_n, it can be obtained by differentiating $T_n(x, y)$ two times with respect to x and setting $(x, y) = (1, 1)$.

Proposition 4.1 *For the CST scheme we have that Var(0) = 0.25, and for $n \geq 1$*

$$\begin{aligned}
\mathit{Var}(n) &= 2^{n-2} + 4^{n-1} - 3\sum_{k=1}^{n} 2^{n-k-2^k} - N\sum_{k=1}^{n}\sum_{l=1}^{k-2} 2^{l-2^{k-l-1}}\\
&\quad + \sum_{k=1}^{n} 2^{n-k+1}\left(\sum_{l=0}^{k-2} 2^{l-2^{k-l-1}}\right)^2 - \left(\frac{N}{2} - \sum_{k=0} 2^{k-N2^{-k}}\right)^2.
\end{aligned}$$

We now sketch a proof for the SD scheme. The only required modification from the CST scheme is the error estimate stated in the following lemma.

Lemma 4.5 *For sufficiently small $\delta > 0$, $|x-1| \le \delta$, $|y-1| \le \delta$, and for all $n \ge 0$, we have*

$$|S_n(x,y)| \ge (4/3)(4/3)^{2^n};$$

for all $n \ge 1$, we have

$$|D_{n-1}(x,y)| \le \frac{4}{9}(4/3)^{2^n}.$$

PROOF. It is easy to check that the lemma holds for $0 \le n \le 3$, as $1 + xy$ is close to 2 when x and y are both near 1.

For $n \ge 4$, we first note, by induction, that

$$|D_{n-1}(x,y)| \le 2\delta \left(n2^n (1+\delta)^{2^n} (4/3)^{2^{n-2}} \right) \le \frac{4}{9}(4/3)^{2^n}.$$

Hence

$$|S_n(x,y)| \ge |S_{n-1}(x,y)|^2 - |D_{n-1}(x,y)| \ge (4/3)(4/3)^{2^n}.$$

■

The following lemma follows immediately from the previous lemma and it is parallel to Lemma 4.3.

Lemma 4.6 *For all $n \ge 0$, $|x-1| \le \delta$, $|y-1| \le \delta$, $x = e^s$ and $y = e^t$, we have,*

$$S_n(e^s, e^t) = \exp\left(2^n u(s,t) + O\left((9/10)^{2^n}\right)\right), \tag{4}$$

where

$$u(s,t) = \ln(1+xy) + \sum_{j \ge 0} 2^{-j-1} \ln\left(1 + D_j(x,y) S_j^{-2}(x,y)\right) \tag{5}$$

is an analytic function in a neighbor of $(s,t) = (0,0)$.

Define

$$c_1(j) = \left.\frac{\partial S_j(x,y)}{\partial x}\right|_{(1,1)}, \; c_2(j) = \left.\frac{\partial S_j(x,y)}{\partial y}\right|_{(1,1)},$$

$$d_1(j) = \left.\frac{\partial D_j(x,y)}{\partial x}\right|_{(1,1)}, \; d_2(j) = \left.\frac{\partial^2 D_j(x,y)}{\partial x^2}\right|_{(1,1)}, \; d_3(j) = \left.\frac{\partial^2 D_j(x,y)}{\partial x \partial y}\right|_{(1,1)}.$$

With some algebra, we may obtain

$$\begin{aligned}
\mu_1 &= \frac{1}{2} + \sum_{j \ge 0} 2^{-j-1-2^{j+1}} d_1(j), \; \mu_2 = \frac{1}{2}, \; \sigma_2^2 = \frac{1}{4}, \\
\sigma_1^2 &= \mu_1 - \frac{1}{4} + \sum_{j \ge 0} 2^{-j-1-2^{j+1}} \left(d_2(j) - 2^{-2^{j+1}} d_1^2(j) - 2^{2-2^j} c_1(j) d_1(j) \right), \\
\sigma_{1,2} &= \frac{1}{4} + \sum_{j \ge 0} 2^{-j-1-2^{j+1}} \left(d_3(j) - 2^{1-2^j} c_2(j) d_1(j) \right).
\end{aligned}$$

The first few values of $c_1(j), c_2(j), d_1(j), d_2(j)$ and $d_3(j)$ can be computed using the following recursions

$$
\begin{aligned}
d_1(0) &= -1,\ d_1(1) = -5,\ d_1(2) = -13,\\
d_1(j) &= -\left(1 + 3\times 2^j + \sum_{i=1}^{j-2} 2^{j-i}\left(2^{2^i}-1\right)^2\right), \quad j \ge 3,\\
d_2(j) &= -2(2^{j+2}-3), \quad 0 \le j \le 2,\\
d_2(j) &= -2\left(1 + 3\times 2^j + \sum_{i=1}^{j-2} 2^{j-i}\left(2^{2^i}-1\right)^2\right) - 4\sum_{i=1}^{j-2} 2^{j-i}\left(2^{2^i}-1\right)(c_1(i)-1), \quad j \ge 3,\\
d_3(j) &= -2^{j+1}(2^{j+2}-j-3), \quad 0 \le j \le 2,\\
d_3(j) &= -2\left(2^j + 2^j\left(3\times 2^j - 2\right)\right.\\
&\quad \left. + \sum_{i=1}^{j-2} 2^{j-i}\left(\left(2^j-2^i\right)\left(2^{2^i}-1\right)^2 + \left(2^{2^i}-1\right)(c_2(i)-2^i)\right)\right), \quad j \ge 3,\\
c_1(0) &= 1,\ c_1(j) = 2^{1+2^{j-1}} c_1(j-1) + d_1(j-1), \quad j \ge 1,\\
c_2(0) &= 1,\ c_2(j) = 2^{1+2^{j-1}} c_2(j-1), \quad j \ge 1.
\end{aligned}
$$

We note again that all the series appearing above converge very rapidly as the general terms decrease doubly exponentially. By taking the first 5 terms (j from 0 to 4), we obtain the following values (correct to 9 decimal places)

$$\mu_1 = 0.2904691622,\ \sigma_1^2 = 0.080396785,\ \sigma_{1,2} = 0.013328463.$$

5 Conclusions

In this paper we proved that the mean number of encryptions for the complete subtree scheme (CST) and the subset-difference scheme (SD) studied by Park and Blake are indeed good estimates for the number of encryptions used by this scheme. We did so by proving a normal limiting distribution for the number of encryptions, as the number of users became large. Indeed, we not only provided the asymptotic normality for the number of encryptions, but also did so for the combined number of encryptions and number of privileged users in a random privileged set. The proofs required a two-dimensional quasi power theorem due to Heuberger. A normal limit distribution also holds for the number of encryptions for the layered subset-difference scheme (LSD), also studied by Park and Blake.

References

[1] AACS: Advanced Access Content System, `http://www.aacsla.com/`.

[2] AACS (Advanced Access Content System), *Introduction and Common Cryptographic Elements*, Rev. 0.91, Feb. 2006, `http://www.aacsla.com/specifications/`.

[3] D. Halevy and A. Shamir, The LSD broadcast encryption scheme, in *CRYPTO 2002*, Lecture Notes in Computer Science **2442** (2002), pp. 47–60.

[4] C. Heuberger, Hwang's quasi-power-theorem in dimension two, *Quaestiones Mathematicae*, **30** (2007), pp. 507–512.

[5] H.K. Hwang, On convergence rates in the central limit theorems for combinatorial structures, *European J. Combin.* **19** (1998), pp. 329–343.

[6] D. Naor, M. Naor and J. Lotspiech, Revocation and tracing schemes for stateless receivers, in *CRYPTO 2001*, Lecture Notes in Computer Science **2139** (2003), pp. 374–391.

[7] E.C. Park and I.F. Blake, On the mean number of encryptions for tree-based broadcast encryption schemes, *Journal of Discrete Algorithms*, **4** (2006), pp. 215–238.

[8] D. Wallner, E. Harder and R. Agee, Key management for multicast: Issues and architectures, Sep. 1998, Internet daft available from `http://www.ietf.org/ID.html`.

[9] C.K. Wong, M. Gouda and S. Lam, Secure group communications using key graphs, in *SIGCOMM 1998*, ACM Press, New York, 1998, pp. 68–79.