# AVERAGE-CASE ANALYSIS OF TWO REVOCATION SCHEMES FOR STATELESS RECEIVERS

CHRISTOPHER EAGLE, MOHAMED OMAR,
DANIEL PANARIO AND BRUCE RICHMOND

Abstract. We consider the mean and variance of the number of encryptions necessary over all subsets of privileged users 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 [4], give the mean number of encryptions for these schemes. We continue their analysis and show that the standard deviations are small compared to the means as the number of users gets large. Therefore, the mean number is a good estimate of the number of necessary encryptions used by these schemes.

*Keywords: Key distribution schemes, the subset-difference scheme (SD), the complete subtree scheme (CST), broadcast encryption.*

## 1. Introduction

We consider the problem of a center broadcasting an encrypted message to a group of users such that some subset is considered revoked and should not be able to obtain the content of the broadcasted message even if all revoked users collaborate. Various encryption schemes have been proposed to solve this problem which arises with 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 who decrypts 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 priveleged (non-revoked user) has at least one decryption key. If the subset of priveleged users is arbitrary and dynamically changing the problem of minimizing user storage and the number of encryptions to ensure system security arises.

To achieve the above goals, several key distribution schemes use a balanced binary tree structure. Two examples that we consider are the subset-difference scheme (SD), introduced by Naor, Naor and Lotspiech [3] and the complete subtree scheme (CST), introduced independently by Wallner, Harder, Agee [5] and Wong, Gouda and Lam [6]. In the CST scheme each user is represented as a unique leaf node in a balanced binary tree. Every node is assigned a key and each user holds the keys which are on the path

*Date*: October 11, 2006.

from its leaf node to its root node. In the SD scheme each user is also represented as a unique leaf node in a balanced binary tree. Instead of assigning a key to every node, the SD scheme assigns a key to every subset difference $S_{i,j} = S_i/S_j$ where node $j$ is a descendent of the node $i$ and $S_i$ is the subtree rooted at the node $i$. If $i = j, S_{ij}$ is empty and no key is assigned. We note that a unique key is assigned to every $S_{ij}$. For a detailed explanation of these schemes see [4].

A practical application of the work of Naor, Naor and Lotspiech is to be found in the forthcoming HD-DVD and BluRay technologies. These two technologies, which are competing to become successor of the DVD technology, each rely on the Advanced Access Content System (AACS) for their 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 subset-difference scheme by AACS each device capable of decoding the HD-DVD or BluRay 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.

In [4], Park and Blake give generating functions that entail the exact mean number of encryptions for several key distribution schemes, including CST and SD. We shall need the results of Park and Blake, and indeed, our analysis can be regarded as a continuation of their work. We provide the standard deviation and variance for the CST and the SD schemes. Since these standard deviations are small with respect to their means, our results imply that the average number of encryptions provided by Park and Blake are indeed good estimates for the number of encryptions in these methods.

The structure of this papers is as follows. In Section 2, we review Park and Blake [4] results. In Section 3, we give the main results of this paper: standard deviation and variance for the number of encryptions for the CST and the SD schemes. Conclusions and further work are given in Section 4.

## 2. Mean Number of Encryptions

We now start from the analysis and notation of Park and Blake [4]. They suppose that there are $N = 2^n$ users in the system. We denote by $(i, j)$-priveleged users a set of $j$ priveleged users that require $i$ encryptions. We note that in the binary tree representation a given set of priveleged users can be partitioned into users in the left subtree and users in the right subtree. The number of $(i, j)$-priveleged users in a system of $2^n$ users can be expressed as the number of $(i^{'}, j^{'})$-priveleged users in the left subtree and $(i-i^{'}, j-j^{'})$-priveleged 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. 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 [4] give a Recurrence for the generating function $T_n(x,y) = \sum_{j=0}^{2^n} \sum_{i=0}^{j} a_{ij}^{(n)} x^i y^j$ for the CST scheme.

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

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

They also provide a recurrence for the generating function $S_n(x,y)$ for the SD scheme, where $S_n(x,y) = \sum_{j=0}^{2^n} \sum_{i=0}^{j} a_{ij}^{(n)} x^i y^j$.

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

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

*where*

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

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

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

Park and Blake use their generating functions to give exact expressions for the mean number of encryptions over all privileged sets. They assume that each of the $2^N$ possible privileged sets have the same probability. 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 Theorem 2.1 and Theorem 2.2, respectively. They prove the following exact mean number estimate.

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

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

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

For the SD scheme they prove the following result.

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

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

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

We take the Park-Blake analysis a bit further for the CST and SD schemes. We need Chebyschev's Inequality.

**Theorem 2.5** (Chebyschev). *If a random variable $X$ has mean $m(n)$ and standard deviation $\sigma(n)$, then the probability that $X$ is at least $t$ standard deviations from the mean satisfies*

$$P\{|X - m(n)| \geq t\sigma(n)\} \leq \frac{1}{t^2}.$$

If $m(n)/\sigma(n) \to 0$ as $n \to \infty$ then $X \sim m(n)$ will be observed with probability tending to one as $n$ tends to $\infty$.

Our efforts in the next section are directed to proving that the standard deviations are small compared to the means. Park and Blake derive asymptotic estimates for the means which they find to be accurate numerical estimates of the means. Our results combined with Chebyschev's Inequality prove that their mean estimates will be accurate estimates for the actual number of encryptions required in these schemes.

## 3. Variance Estimates

We now include a proof of Theorem 2.3 due to Park and Blake. They leave this proof to the reader and indeed it is possible to get the required exact result by asking MAPLE. However, we think it is helpful to see the derivation because we shall derive more complicated results using similar ideas.

Proof (Theorem 2.3). From the definition of $m(n)$ in (1), we have

$$m_{\mathsf{CST}}(n) = \frac{1}{2^N}\frac{\partial T_n(x,y)}{\partial x}(1,1),$$

where $T_0(x,y) = 1 + xy$, and for $n \geq 1$ we have $T_n(x,y) = T_{n-1}(x,y)^2 + (1-x)xy^{2^n}$. So

$$m_{\mathsf{CST}}(0) = \frac{1}{2^N}\frac{\partial(1+xy)}{\partial x}(1,1) = 0.5.$$

For $n \geq 1$, and using that $N = 2^n$, we have

$$\begin{aligned}\frac{\partial T_n(x,y)}{\partial x}(1,1) &= \left(2T_{n-1}(x,y)\frac{\partial T_{n-1}(x,y)}{\partial x} + y^N - 2xy^N\right)(1,1)\\ &= 2T_{n-1}(1,1)\frac{\partial T_{n-1}(x,y)}{\partial x}(1,1) - 1.\end{aligned}$$

Now $T_0(1,1) = 2$, and for $r \geq 1, T_r(1,1) = T_{r-1}(1,1)^2 + (1-1)1 = T_{r-1}(1,1)^2$. Thus, by induction, $T_r(1,1) = 2^{2^r}$ for all $r \geq 0$. If we let

$$F(n) = \frac{\partial T_n(x,y)}{\partial x}(1,1),$$

we have

$$F(n) = 2T_{n-1}(1,1)F(n-1) - 1 = (2)2^{2^{n-1}}F(n-1) - 1 = 2^{2^{n-1}+1}F(n-1) - 1.$$

Thus, for the mean number $m_{\mathtt{CST}}(n)$, we have

$$\begin{aligned} (2) \qquad m_{\mathtt{CST}}(n) &= \frac{1}{2^{2^n}}F(n) = 2^{-2^n}2^{2^{n-1}+1}F(n-1) - \frac{1}{2^N} \\ &= 22^{2^{n-1}-2^n}F(n-1) - 2^{-N} = 2m_{\mathtt{CST}}(n-1) - 2^{-N}. \end{aligned}$$

We now proceed by induction. We have seen that $m_{\mathtt{CST}}(0) = 0.5$, and $m_{\mathtt{CST}}(1) = 2m_{\mathtt{CST}}(0) - 2^{-2^1} = 1 - 1/4 = 3/4$. We rewrite this as

$$\frac{2}{2} - \left(\sum_{k=0}^{0} 2^{k-22^{-k}}\right) = \frac{N}{2} - \sum_{k=0}^{0} 2^{k-N2^{-k}} \text{ since } N = 2 \text{ for } n = 1.$$

The theorem is true for $n = 1$. Assume the theorem holds for $n - 1 \geq 1$. Then, by (2), $m_{\mathtt{CST}}(n) = 2m_{\mathtt{CST}}(n-1) - 2^{-2^n}$ so

$$\begin{aligned} m_{\mathtt{CST}}(n) &= 2\left(\frac{2^{n-1}}{2} - \sum_{k=0}^{n-2} 2^{k-2^{n-1}2^{-k}}\right) - 2^{-2^n} \\ &= 2^{n-1} - 2\sum_{k=0}^{n-2} 2^{k-2^{n-1}2^{-k}} - 2^{-2^n} \\ &= \frac{N}{2} - \sum_{k=0}^{n-2} 2^{(k+1)-2^{n-1}2^{-k}} - 2^{-2^n} \\ &= \frac{N}{2} - \sum_{l=1}^{n-1} 2^{l-2^{n-1}2^{-l+1}} - 2^{-2^n} \\ &= \frac{N}{2} - \left(\sum_{l=1}^{n-1} 2^{l-2^n2^{-l}} + 2^{-2^n}\right) = \frac{N}{2} - \sum_{l=0}^{n-1} 2^{l-N2^{-l}}. \end{aligned}$$

Therefore, the theorem holds for $m_{\mathtt{CST}}(n)$. ■

3.1. **The Complete Subtree Scheme.** We now turn to the variance for the CST scheme. In this section, for simplicity, we use $m(n)$ instead of $m_{\mathtt{CST}}(n)$. We define the second moment $\mu(n)$ by

$$\mu(n) = \frac{1}{2^N}\sum_i \sum_j i^2 a_{ij}^{(n)},$$

and the variance by $\mathrm{Var}(n) = \mu(n) - m(n)^2$. We observe that

$$2^N\mu(n) = \sum_i \sum_j i^2 a_{ij}^{(n)} = \frac{\partial T_n(x,y)}{\partial x}(1,1) + \frac{\partial^2 T_n(x,y)}{\partial x^2}(1,1).$$

Our first result is the following.

**Theorem 3.1.** *For the CST scheme we have that Var(0) = 0.25 and for* $n \geq 1$

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

Proof. We recall from Theorem 2.1 that $T_n(x,y) = T_{n-1}(x,y)^2 + (1-x)xy^N$ and $T_0(x,y) = 1 + xy$. For $n = 0$ we find

$$\frac{\partial T_0(x,y)}{\partial x}(1,1) = \frac{\partial(1+xy)}{\partial x}(1,1) = y(1,1) = 1,$$

and so from our formula for $\mu(0)$ in terms of partial derivatives we have

$$\mu(0) = \frac{1}{2^{2^n}}(1+0) = 1/2.$$

Thus, the theorem holds for $\mathrm{Var}(0) = 1/2 - (1/2)^2 = 1/4$. For $n \geq 1$

$$\begin{aligned} \frac{\partial T_n(x,y)}{\partial x} &= 2T_{n-1}(x,y)\frac{\partial T_{n-1}(x,y)}{\partial x} + y^N - 2xy^N \\ \frac{\partial^2 T_n(x,y)}{\partial x^2} &= 2\left[\left(\frac{\partial T_{n-1}(x,y)}{\partial x}\right)^2 + \frac{\partial x^2 T_{n-1}(x,y)}{\partial^2} T_{n-1}(x,y)\right] - 2y^N. \end{aligned}$$

So

$$\begin{aligned} \frac{\partial T_n(x,y)}{\partial x}(1,1) &= 2T_{n-1}(1,1)\frac{\partial T_{n-1}(x,y)}{\partial x}(1,1) + 1 - 2 \\ &= 22^{2^{n-1}}\frac{\partial T_{n-1}(x,y)}{\partial x}(1,1) - 1 \end{aligned}$$

since $T_k(x,y)(1,1) = 2^{2^k}$ for each $k \geq 0$. Furthermore, we have

$$\begin{aligned} &\frac{\partial^2 T_{n-1}(x,y)}{\partial x^2}(1,1) \\ &= 2\left[\left(\frac{\partial T_{n-1}(x,y)}{\partial x}(1,1)\right)^2 + \frac{\partial^2 T_{n-1}(x,y)}{\partial x^2}(1,1)T_{n-1}(1,1) - 1\right] \\ &= 2\left(\frac{\partial T_{n-1}(x,y)}{\partial x}(1,1)\right)^2 + (2)2^{2^{n-1}}\frac{\partial^2 T_{n-1}(x,y)}{\partial x^2}(1,1) - 2. \end{aligned}$$

So, using again the partial derivatives expression for $\mu(n)$, we get

$$\begin{aligned}
\mu(n) &= \frac{1}{2^{2^n}}\left(\frac{\partial T_n(x,y)}{\partial x}(1,1)+\frac{\partial^2 T_n(x,y)}{\partial x^2}(1,1)\right)\\
&= 2\frac{2^{2^{n-1}}}{2^{2^n}}\frac{\partial T_{n-1}(x,y)}{\partial x}(1,1)-\frac{1}{2^{2^n}}\\
&\quad +2\frac{1}{2^{2^n}}\left(\frac{\partial T_{n-1}(x,y)}{\partial x}(1,1)\right)^2+2\frac{2^{2^{n-1}}}{2^{2^n}}\frac{\partial^2 T_{n-1}(x,y)}{\partial x^2}(1,1)-\frac{2}{2^{2^n}}\\
&= 2\left(\frac{1}{2^{2^{n-1}}}\frac{\partial T_{n-1}(x,y)}{\partial x}(1,1)+\frac{1}{2^{2^{n-1}}}\frac{\partial^2 T_{n-1}(x,y)}{\partial x^2}(1,1)\right)\\
&\quad +2\left(\frac{1}{2^{2^{n-1}}}\frac{\partial T_{n-1}(x,y)}{\partial x}(1,1)\right)^2-\frac{3}{2^{2^n}}\\
&= 2\mu(n-1)+2m(n-1)^2-\frac{3}{2^{2^n}}.
\end{aligned}$$

So

$$\frac{1}{2^{2^n}}\sum_i\sum_j i^2 a_{ij}^{(n)}=\mu(n)=2\mu(n-1)+2m(n-1)^2-\frac{3}{2^N}. \tag{3}$$

Now, from Theorem 2.3, we have

$$m(n-1)=\frac{2^{n-1}}{2}-\sum_{k=0}^{n-2}2^{k-2^{n-1}2^{-k}}=2^{n-2}-\sum_{k=0}^{n-2}2^{k-2^{n-k-1}}.$$

We can put our explicit formula into (3) and solve the resulting first order difference equation (or ask MAPLE), to obtain

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

We finally get

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

■

For example with $\mathrm{Std}(n) = \sqrt{\mathrm{Var}(n)}$ this yields $m(0) = 0.5, \mathrm{Std}(0) = 0.5$, $m(1) = 0.75, \mathrm{Std}(1) = \sqrt{3}/4$, $m(2) = 23/16, \mathrm{Std}(2) = \sqrt{95}/16$, $m(15) = 11759.96175\cdots, \mathrm{Std}(15) = 55.53559796\cdots$, and so on.

We note that the numerical results suggest that the standard deviation is quite small compared to the mean. We next prove that this is indeed the case for the CST scheme. We will later show that this is also the case for the SD scheme.

**Theorem 3.2.** *Let $m(n)$ and $Std(n)$ denote the mean and standard deviation for the CST scheme with $2^n$ users. Then*

$$\frac{Std(n)}{m(n)} \leq \frac{1}{2^{(n-5)/2}}.$$

*Furthermore, we have with probability tending to 1*

$$\sum_{j\geq 0} a_{ij}^{(n)} = m(n) + O\left(t(n)Std(n)\right),$$

*where $t$ is any function of $n$ that goes to infinity with $n$ and $0 \leq i \leq 2^n$.*

Proof. We have from Theorem 2.3

$$\begin{aligned} m(n) &= 2^{n-1} - \sum_{k=0}^{n-1} 2^{k-2^{n-k}} \geq 2^{n-1} - \sum_{k=0}^{n-1} 2^{n-3-3k} \\ &= 2^{n-1} - 2^{n-3}\sum_{k=0}^{n-1} 8^{-k} \geq 2^{n-1} - 2^{n-3}\frac{8}{7} \geq 2^{n-2}. \end{aligned}$$

Thus $m(n)^2 \geq 2^{2n-4}$. Moreover, we have

$$m(n) = 2^{n-1} - \sum_{k=0}^{n-1} 2^{k-2^{n-k}} \leq 2^{n-1}.$$

Using (3) and (2), we have

$$\begin{aligned} \mathrm{Var}(n) &= 2\mathrm{Var}(n-1) + 4m(n-1)^2 - m(n)^2 - \frac{3}{2^N} \\ &= 2\mathrm{Var}(n-1) + \frac{4m(n-1)-3}{2^N} - \frac{1}{2^{2N}} \\ &\leq 2\mathrm{Var}(n-1) + \frac{42^{n-2}}{2^N} - \frac{3}{2^N} - \frac{1}{2^{2N}} \\ &\leq 2\mathrm{Var}(n-1) + \frac{N}{2^N} \leq 2\mathrm{Var}(n-1) + 1. \end{aligned}$$

Thus, for all $n$, $\mathrm{Var}(n) \leq t_n$ where $t_n = 2t_{n-1} + 1$ and $t_0 = \mathrm{Var}(0) = 1/4$. Solving for $t_n$, we find that $t_n = (5)2^{n-2} - 1$. Thus, $\mathrm{Var}(n) \leq (5)2^{n-2} - 1 \leq (8)2^{n-2} = 2^{n+1}$. Hence, we can conclude that

$$\frac{\mathrm{Var}(n)}{m^2(n)} \leq \frac{2^{n+1}}{2^{2n-4}} = \frac{1}{2^{n-5}}.$$

The second part of the theorem follows immediately from Theorem 2.5. ■

3.2. **The Subset-Difference Scheme.** We now consider the SD scheme. We shall be content to give a recurrence to compute variance and show that the standard deviation is very small compared to the mean. The recurrence is more useful to compute the variances than any explicit formula we have found. Again for the SD scheme we use $m(n)$ instead of $m_{\mathsf{SD}}(n)$, and we have $\mathrm{Var}(n) = \mu(n) - m(n)^2$. We have, with the notation of Theorem 3.1

$$\mu(n) = \frac{1}{2^N} \sum_i \sum_j i^2 a_{ij}^{(n)} = \frac{1}{2^N} \left( \frac{\partial^2 S_n(x,y)}{\partial x^2 x}(1,1) + \frac{\partial S_n(x,y)}{\partial x}(1,1) \right),$$

where

$$\begin{aligned}
S_0(x,y) &= 1 + xy, \\
S_n(x,y) &= S_{n-1}(x,y)^2 + D_{n-1}(x,y) \qquad \text{for} \quad n \geq 1, \\
D_0(x,y) &= (1-x)xy^2, \\
D_{n-1}(x,y) &= (1-x)x \left[ y^{2^n} + 2^n y^{2^n} \sum_{i=0}^{n-2} 2^{-i} y^{-2^i} \right] \qquad \text{for} \quad 2 \leq n \leq 3,
\end{aligned}$$

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

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

We begin with the initial cases $n = 0, 1, 2, 3$. For $n = 0$ we have

$$\begin{aligned}
m(0) &= \frac{1}{2^{2^0}} \frac{\partial S_0(x,y)}{\partial x}(1,1) = \frac{1}{2}, \\
\mu(0) &= \frac{1}{2^{2^0}} \left( \frac{\partial S_0(x,y)}{\partial x}(1,1) + \frac{\partial^2 S_0(x,y)}{\partial x^2}(1,1) \right) = \frac{1}{2},
\end{aligned}$$

thus $\mathrm{Var}(0) = 1/2 - (1/2)^2 = 1/4$.

For $n = 1$ we have

$$\begin{aligned}
m(1) &= \frac{1}{2^{2^1}} \frac{\partial S_1(x,y)}{\partial x}(1,1) = \frac{1}{4} \frac{\partial}{\partial x} \left( (1+xy)^2 + (1-x)xy^2 \right)(1,1) = \frac{3}{4}, \\
\mu(1) &= \frac{1}{4} \frac{\partial^2 S_1(x,y)}{\partial x^2}(1,1) + m(1) = 3/4,
\end{aligned}$$

therefore in this case we have $\mathrm{Var}(1) = 3/4 - (3/4)^2 = 3/16$.

In the same way we compute the values for $n = 2$ and $n = 3$, summarized in Table 1.

| $n$ | $m(n)$ | $\mu(n)$ | $\text{Var}(n)$ |
|---|---|---|---|
| 0 | 0.5 | 0.5 | 0.25 |
| 1 | 0.75 | 0.75 | 0.1875 |
| 2 | 1.1875 | 1.6875 | 0.2773 |
| 3 | 2.3242 | 6.0429 | 0.6409 |

Table 1. Values of $m(n)$, $\mu(n)$ and $\text{Var}(n)$ for the SD scheme, $n = 0, 1, 2$ and 3.

We now come to the main case, $n \geq 4$. We defined in Theorem 2.2

$$\begin{aligned} D_{n-1}(x,y) &= (1-x)xy^N \left(1 + N\sum_{i=0}^{1} 2^{-i}y^{-2^i} \right. \\ &\left. + \frac{N}{2}\sum_{i=1}^{n-3} 2^{-i}y^{-2^{i+1}} \left(S_i(x,y) - xy^{2^i}\right)^2\right). \end{aligned}$$

Let $\phi_n(x,y)$ be defined by

$$\phi_n(x,y) = 1 + N\sum_{i=0}^{1} 2^{-i}y^{-2^i} + \frac{N}{2}\sum_{i=1}^{n-3} 2^{-i}y^{-2^{i+1}} \left(S_i(x,y) - xy^{2^i}\right)^2. \tag{4}$$

We now derive recurrences for the mean and standard deviation with the SD scheme. From Theorem 2.2 we have

$$S_n(x,y) = S_{n-1}(x,y)^2 + y^N\left(x - x^2\right)\phi_n(x,y).$$

Thus,

$$\begin{aligned} \frac{\partial S_n(x,y)}{\partial x} &= 2S_{n-1}(x,y)\frac{\partial S_{n-1}(x,y)}{\partial x} \\ &+ y^N\left((1-2x)\phi_n(x,y) + \left(x - x^2\right)\frac{\partial \phi_n(x,y)}{\partial x}\right). \end{aligned}$$

Next we observe that $S_n(1,1) = 2^{2^n}$. Indeed, recalling from Theorem 2.2 the definition of $S_n(x,y)$, we find that $S_0(1,1) = 1 + 1 * 1 = 2$. Considering $D_n(x,y)$ for any $n \geq 0$, we observe that there is a multiplying factor $1-x$. So we have $D_n(1,1) = 0$. Thus, for $n \geq 1$, $S_n(1,1) = S_{n-1}(1,1)^2 + 0$. Solving this recurrence we obtain that for all $n \geq 0$, $S_n(1,1) = 2^{2^n}$.

Hence, returning to the expression of the derivative of $S_n(x,y)$, we have

$$\begin{aligned} \frac{\partial S_n(x,y)}{\partial x}(1,1) &= 2S_{n-1}(1,1)\frac{\partial S_{n-1}(x,y)}{\partial x}(1,1) + (1-2)\phi_n(1,1) \\ &= (2)2^{2^{n-1}}\frac{\partial S_{n-1}(x,y)}{\partial x}(1,1) - \phi_n(1,1). \end{aligned}$$

Furthermore we have

$$\begin{aligned} m(n) &= \frac{1}{2^{2^n}}\frac{\partial S_n(x,y)}{\partial x}(1,1) = \frac{(2)2^{2^{n-1}}}{2^{2^n}}\frac{\partial S_n(x,y)}{\partial x}(1,1) - \frac{\phi_n(1,1)}{2^{2^n}} \\ &= 2m(n-1) - \frac{\phi_n(1,1)}{2^{2^n}} \\ \frac{\partial^2 S_n(x,y)}{\partial x^2} &= 2\left(\left(\frac{\partial S_{n-1}(x,y)}{\partial x}\right)^2 + S_{n-1}(x,y)\frac{\partial^2 S_{n-1}(x,y)}{\partial x^2}\right) \\ &\quad + y^N\left(-2\phi_n(x,y) + (1-2x)\frac{\partial \phi_n(x,y)}{\partial x}\right. \\ &\quad \left. + (1-2x)\frac{\partial \phi_n(x,y)}{\partial x} + \left(x-x^2\right)\frac{\partial^2 \phi_n(x,y)}{\partial x^2}\right). \end{aligned}$$

So we get

$$\begin{aligned} &\frac{\partial^2 S_n(x,y)}{\partial x^2}(1,1) \\ = \;& 2\left(\left(\frac{\partial S_{n-1}(x,y)}{\partial x}(1,1)\right)^2 + S_{n-1}(1,1)\frac{\partial^2 S_{n-1}(x,y)}{\partial x^2}(1,1)\right) \\ & -2\phi_n(1,1) - 2\frac{\partial \phi_n(x,y)}{\partial x}(1,1) + 0 \\ = \;& 2\left(\left(\frac{\partial S_{n-1}(x,y)}{\partial x}(1,1)\right)^2(1,1) + 2^{2^{n-1}}\frac{\partial^2 S_{n-1}(x,y)}{\partial x^2}(1,1)\right) \\ & -2\phi_n(1,1) - 2\frac{\partial \phi_n(x,y)}{\partial x}(1,1). \end{aligned}$$

We may now derive a recurrence for $\mu(n)$

$$\begin{aligned} &\mu(n) \\ = \;& \frac{1}{2^{2^n}}\left(\frac{\partial^2 S_n(x,y)}{\partial x^2}(1,1) + \frac{\partial S_n(x,y)}{\partial x}(1,1)\right) \\ = \;& \frac{1}{2^{2^n}}\left(2\left(\frac{\partial S_{n-1}(x,y)}{\partial x}(1,1)\right)^2 + (2)2^{2^{n-1}}\frac{\partial^2 S_{n-1}(x,y)}{\partial x^2}(1,1)\right. \\ & -2\phi_n(1,1) - 2\frac{\partial \phi_n(x,y)}{\partial x}(1,1) \\ & \left. + (2)2^{2^{n-1}}\frac{\partial S_{n-1}(x,y)}{\partial x}(1,1) - \phi_n(1,1)\right) \\ = \;& \frac{2}{2^{2^n}}\left(\frac{\partial S_{n-1}(x,y)}{\partial x}(1,1)\right)^2 + 2\left(\frac{1}{2^{2^{n-1}}}\frac{\partial^2 S_{n-1}(x,y)}{\partial x^2}(1,1)\right. \\ & \left. + \frac{1}{2^{2^{n-1}}}\frac{\partial S_{n-1}(x,y)}{\partial x}(1,1)\right) - \frac{3}{2^{2^n}}\phi_n(1,1) - \frac{2}{2^{2^n}}\frac{\partial \phi_n(x,y)}{\partial x}(1,1) \end{aligned}$$

$$
\begin{aligned}
&= 2\left(\frac{1}{2^{2^{n-1}}}\frac{\partial S_{n-1}(x,y)}{\partial x}(1,1)\right)^2 + 2\mu(n-1) - \frac{3}{2^{2^n}}\phi_n(1,1) \\
&\quad -\frac{2}{2^{2^n}}\frac{\partial \phi_n(x,y)}{\partial x}(1,1).
\end{aligned}
$$

Finally we have the following form of the recursion

$$
\mu(n) = 2m(n-1)^2 + 2\mu(n-1) - \frac{1}{2^{2^n}}\left(3\phi_n(1,1) + 2\frac{\partial \phi_n(x,y)}{\partial x}(1,1)\right).
$$

For the variance we find

$$
\begin{aligned}
\mathrm{Var}(n) &= \mu(n) - m(n)^2 \\
&= 2m(n-1)^2 + 2\mu(n-1) - \frac{1}{2^{2^n}}\left(3\phi_n(1,1) + 2\frac{\partial \phi_n(x,y)}{\partial x}\right)(1,1) \\
&\quad -\left(2m(n-1) - \frac{\phi_n(1,1)}{2^{2^n}}\right)^2 \\
&= 2\mu(n-1) - 2m(n-1)^2 - \frac{\phi_n^2(1,1)}{2^{2^{n+1}}} \\
&\quad -\frac{1}{2^{2^n}}\left(3\phi_n(1,1) + 2\frac{\partial \phi_n(x,y)}{\partial x}(1,1) - 4m(n-1)\phi_n(1,1)\right).
\end{aligned}
$$

Finally we have

$$
\begin{aligned}
\mathrm{Var}(n) &= 2\mathrm{Var}(n-1) - \frac{\phi_n(1,1)^2}{2^{2^{n+1}}} \\
&\quad -\frac{1}{2^{2^n}}\left(3\phi_n(1,1) + 2\frac{\partial \phi_n(x,y)}{\partial x}(1,1) - 4m(n-1)\phi_n(1,1)\right). \qquad (5)
\end{aligned}
$$

We now show that the standard deviation is very small compared to the mean for the SD sheme also.

**Theorem 3.3.** *The ratio of the standard devation divided by the mean for the SD scheme satisfies, when $n \geq 6$,*

$$
\frac{Std(n)}{m(n)} \leq \frac{4}{2^{n/2}}.
$$

*For the SD scheme we also have with probability tending to* 1

$$
\sum_{j\geq 0} a_{ij}^{(n)} = m(n) + O\left(t(n)Std(n)\right),
$$

*where $t$ is any function of $n$ that goes to infinity with $n$ and $0 \leq i \leq 2^n$.*

Proof. From (5) and Theorem 2.4 we have

$$
\begin{aligned}
\mathrm{Var}(n) &\leq 2\mathrm{Var}(n-1) + \frac{4m(n-1)\phi_n(1,1)}{2^{2^n}} \\
&\leq 2\mathrm{Var}(n-1) + \frac{2^n\phi_n(1,1)}{2^{2^n}}.
\end{aligned}
$$

Furthermore

$$\frac{\phi_n(1,1)}{2^{2^n}} = \frac{1}{2^{2^n}}\left(1+(3)2^{n-1}+2^{n-1}\sum_{i=1}^{n-3}\frac{\left(2^{2^i}-1\right)^2}{2^i}\right)$$

$$\leq \frac{1}{2^{2^n}}\left(1+(3)2^{n-1}+2^{n-1}\sum_{i=1}^{n-3}\frac{2^{2^{i+1}}}{2^i}\right).$$

Now

$$\begin{aligned}\sum_{i=1}^{n-3}2^{2^{i+1}-i} &= 2^{2^{n-2}-(n-3)}+2^{2^{n-3}-(n-4)}+\cdots+2^3\\ &\leq 2^{2^{n-2}-(n-3)}+2^{2^{n-2}-(n-4)}+\cdots+2^{2^{n-2}-3}\leq 2^{2^{n-2}-(n-4)}.\end{aligned}$$

So

$$\begin{aligned}\frac{\phi_n(1,1)}{2^{2^n}} &\leq \frac{1}{2^{2^n}}\left(1+(3)2^{n-1}\right)+\frac{2^{n-1}}{2^{2^n}}\frac{2^{n-2}}{2^{n-4}}\\ &= \frac{1}{2^{2^n}}\left(1+(3)2^{n-1}\right)+8\frac{2^{2^{n-2}}}{2^{2^n}}\\ &= \frac{1}{2^{2^n}}\left(1+(3)2^{n-1}\right)+(8)2^{2^{n-2}(1-2^2)}\\ &= \frac{1}{2^{2^n}}\left(1+(3)2^{n-1}\right)+2^{-2^{n-2}}\\ &\leq \frac{1}{2^{2^{n-2}}}\left(2+2^{n+1}\right)\leq\frac{2^{n+2}}{2^{2^{n-2}}}.\end{aligned}$$

Hence, we have for $n \geq 6$

$$\frac{2^n\phi_n(1,1)}{2^{2^n}} \leq \frac{2^{2n+2}}{2^{2^{n-2}}} \leq 1.$$

In this case

$$\mathrm{Var}(n) \leq 2\mathrm{Var}(n-1)+1 \leq \frac{107543}{524286}2^n - 1 \leq 2^{n-2}.$$

Also we have $m(n) = 2m(n-1) - \frac{\phi_n(1,1)}{2^{2^n}}$ so $m(n) \geq 2m(n-1) - 1 \geq \frac{339}{2698}2^n + 1$ and so $m(n) \geq \frac{1}{8}2^n = 2^{n-3}$. Thus $m(n)^2 \geq 2^{2n-6}$. It follows that

$$\frac{\mathrm{Var}(n)}{m(n)^2} \leq \frac{2^{n-2}}{2^{2n-6}} = 2^{4-n}.$$

The theorem now follows using Theorem 2.5. ■

## 4. Conclusions

In this paper we prove 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 encryption used by these schemes. We did so by providing estimates for the variance and the standard deviation that show them small as compared with the mean number of encryptions, as the number of users gets large.

Park and Blake also give the mean number of encryption for other tree-based broadcast encryption scheme: the *layered subset-difference scheme* (LSD). The recurrences for the LSD scheme are much more involved and we have not been able to show that the standard deviation for this scheme is small compared to the mean. Very likely it is however we must leave that for a future investigation.

## References

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

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

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

[4] 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.

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

[6] 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.

Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
*E-mail address*: `cjeagle@engmail.uwaterloo.ca`

Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
*E-mail address*: `momar@student.math.uwaterloo.ca`

School of Mathematics and Statistics, Carleton University, Ottawa, Ontario K1S 5B6, Canada
*E-mail address*: `daniel@math.carleton.ca`

Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
*E-mail address*: `lbrichmond@math.uwaterloo.ca`