Average-Case Analysis of Revocation Schemes for Stateless Receivers

Daniel Panario
School of Mathematics and Statistics
Carleton University
daniel@math.carleton.ca

Joint work with
C. Eagle, Z. Gao, M. Omar and B. Richmond

Analysis of Algorithms, April 2008

Outline

- Introduction:
  - the problem,
  - applications.
- Key distribution schemes:
  - Complete subtree scheme (CST);
  - Subset difference scheme (SD);
  - Layered subset difference scheme (LSD).
- Generating functions.
- Statistical results.

The problem

A center broadcasts an encrypted message to a group of users:

- some users may not be authorized (revoked users);
- revoked users may collaborate but should not be able to obtain the message;
- revoked users are not fixed (change dynamically);
- encrypting messages can be done multiple times;
- decrypting keys cannot be changed (stateless receivers).

The problem: minimize user storage and number of encryptions, while still ensuring system security.

Applications

• Pay-TV: users are subscribers; users are revoked if they don't pay fee for particular channel.

• DVD movies: users are DVD players, revoked if they are tied to illegal activity;

• Blu-ray technology: security features use subset-difference scheme;

• satellite communications, real-time information update, media content protection, etc.

Complete Subtree Scheme

The complete subtree scheme (CST) is due to Wallner, Harder and Agee (1998) and independently Wong, Gouda and Lam (1998):

- 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 from its leaf node to its root node.

Other binary balanced trees key distribution schemes are: subset difference scheme (SD) and layered subset difference scheme (LSD).

Subset difference scheme

Subset difference scheme (SD): Naor, Naor and Lotspiech, 2003.

SD scheme: each user is represented as a unique leaf node in a balanced binary tree but in the SD scheme a key is assigned to every subset difference $S_{i,j} = S_i / S_j$ where node j is a descendent of 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.

Layered subset difference scheme

Layered subset difference scheme (LSD): Halevy and Shamir, 2002.

LSD scheme: key storage is reduced using layers. A layer is the set of levels between two consecutive multiples of $\log N = n$ levels, where N is number of leaves in the balanced binary tree.

In the LSD scheme, S_{ij} is said to be useful if i is a special level or i and j belong to the same layer. We have that any subset difference S_{ij} is a union of two useful sets $S_{ik} \cup S_{kj}$, for nodes i, k and j. Therefore, one only needs to store the useful sets on the same path saving key storage.

Notation

Park and Blake (2006) assume 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.

The number of (i, j)-priveleged users in a system of 2^n users is 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. We have

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

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 give recurrences for the generating functions of the numbers $a_{ij}^{(n)}$ in the CST, SD and LSD schemes.

Generating functions (CST)

Park and Blake gave generating functions for the CST, SD and LSD schemes.

Theorem. The generating function for the CST scheme is

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

Generating functions (SD)

Theorem. The generating function for the SD scheme is

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

where

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

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

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

Generating functions (LSD)

Theorem. The generating function for the LSD scheme is

$$L_n(x,y) = H_n^n(x,y),$$

where

(1) If $0 \leq q \leq \sqrt{n}$, $H_q^n(x,y) = S_q(x,y)$ where $S_q(x,y)$ is the generating function for the SD scheme for 2^q users.

(2) If $q = k\sqrt{n}$ for some integer k,

$$\begin{aligned} H_q^n(x,y) &= H_{q-1}^n(x,y)^2 + (1-x)xy^{2^q} \\ &+(1-x)xy^{2^q}2^q \sum_{q-\sqrt{n}}^{q-2} 2^{-i}y^{-2^i}\left(H_{i-1}^n(x,y) - xy^{2^{i-1}}\right)^2 \\ &+(1-x^2)xy^{2^q}2^q \sum_{2}^{q-\sqrt{n}-1} 2^{-i}y^{-2^i}\left(H_{i-1}^n(x,y) - xy^{2^{i-1}}\right)^2 \\ &+(1-x^2)xy^{2^q}2^q \sum_{i=0}^{1} 2^{-i}y^{-2^i}. \end{aligned}$$

(3) If $q = 1 + k\sqrt{n}$ for some integer k,

$$H_q^n(x,y) = H_{q-1}^n(x,y)^2 + (1-x)xy^{2^q}.$$

(4) If $q = 2 + k\sqrt{n}$ for some integer k,

$$\begin{aligned} H_q^n(x,y) &= H_{q-1}^n(x,y)^2 + (1-x)xy^{2^q} \\ &\quad +4(1-x)xy^{2^q-2^{q-2}}(H_{q-2}^n(x,y) - xy^{2^{q-2}}). \end{aligned}$$

(5) For all other cases,

$$\begin{aligned} H_q^n(x,y) &= H_{q-1}^n(x,y)^2 + (1-x)xy^{2^q} \\ &\quad +(1-x)xy^{2^q}2^q \sum_{i=s(q)+1}^{q-2} 2^{-i}y^{-2^i}\left(H_{i-1}^n(x,y) - xy^{2^{i-1}}\right)^2 \\ &\quad +(1-x)xy^{2^q-2^{s(q)}}2^{q-s(q)}\left(H_{s(q)}^n(x,y) - xy^{2^{s(q)}}\right) \\ &\quad +(1-x^2)xy^{2^q}2^q \sum_{i=0}^{1} 2^{-i}y^{-2^i}, \end{aligned}$$

where $s(q) = \lfloor q/\sqrt{n} \rfloor \sqrt{n}$ refers to the highest special level in a balanced subtree for 2^q users.

Mean number of encryptions

Park and Blake use the above generating functions to give exact expressions for the mean number of encryptions over all privileged sets for the three considered schemes. They assume that each of the 2^N possible privileged sets have the same probability. The mean number of encryption is defined by

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

where $G_n(x,y)$ can be either $T_n(x,y)$, $S_n(x,y)$ or $L_n(x,y)$, as defined before.

They prove the following exact mean number estimates.

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

The mean number of encryptions over all privileged sets for the SD scheme is given by, for $n \geq 4$,

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

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

The mean number of encryptions over all privileged sets for the LSD scheme is given by

$$m_{\mathsf{LSD}}(n) = \frac{N}{2\sqrt{n}} m_{\mathsf{SD}}(\sqrt{n}) + \sum_{i=0}^{\sqrt{n}-2} 2^{\sqrt{n}i} C_{\sqrt{n}-i},\ n \geq 16,$$

where $m_{\mathsf{SD}}(\sqrt{n})$ is the mean number of encryptions over all privileged sets for the SD scheme with $2^{\sqrt{n}}$ users, $A = 2^{\sqrt{n}}$ and

$$\begin{aligned} C_k \quad = \quad & -2^{2A^{k-1}-1} A - 2^{-3A^{k-1}} A + 3\left(2^{-4A^{k-1}-2} A\right) - \sum_{i=1}^{\sqrt{n}-3} 2^{-\frac{A^k}{2^i}+i} \\ & - \sum_{i=0}^{\sqrt{n}-3} 2^{-A^k 2^{-i}} A^k \sum_{j=(k-1)\sqrt{n}+1}^{k\sqrt{n}-2-i} 2^{-j}\left(2^{2^j} - 2^{2^{j-1}+1} + 1\right) \\ & - A 2^{-A^k}\left(2^{A^{k-1}} - 2^{A^{\frac{k-1}{2}}+1} + 1\right) - \sum_{i=1}^{\sqrt{n}-3} 2^{-\frac{A^k}{2^i}} A\left(2^{A^{k-1}} - 1\right) \\ & - 2A^k 2^{-A^k} \sum_{j=2}^{(k-1)\sqrt{n}-1} 2^{-j}\left(2^{2^j} - 2^{2^{j-1}+1} + 1\right) 2^{-A^k} - 3\left(A^k 2^{-A^k}\right). \end{aligned}$$

We take the Park-Blake analysis a bit further by providing limiting distributions for the number of encryptions for these schemes.

In a similar way to Park and Blake paper, one can prove results like:

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

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

But it is hard to extend these results beyond the second moment.

We require Hwang *quasi-power theorem* that give a central limit theorem and convergence rate for a sequence of random variables with moment generating function obeying a quasi-power form

Theorem (Hwang). Let $\{X_n\}_{n\geq 1}$ be a sequence of integral random variables. Assume that the moment generating function asymptotically satisfies

$$M_n(s) = \sum_{m\geq 0} \mathbb{P}(X_n = m)e^{ms} = e^{(u(s)\phi(n)+v(s))}(1 + O(1/\alpha_n)),$$

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

(1) $u(s)$ and $v(s)$ are analytic for $|s| \leq \tau$ and independent of n; and $u''(0) \neq 0$;

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

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

$$\mathbb{P}\left(\frac{X_n - u'(0)\phi(n)}{\sqrt{u''(0)\phi(n)}} < x\right) = \Phi(x) + O\left(\frac{1}{\sqrt{\phi(n)}} + \frac{1}{\alpha_n}\right),$$

uniformly with respect to $x \in \mathbb{R}$, where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2} dy.$$

In our problem we have two sequences of random variables (the number of encryptions and privileged users in a random privileged set). Thus, we require a bivariate version of the quasi-power theorem to deal with the joint distribution. We use Heuberger (2007) extension to two dimensions.

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

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

Theorem (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.,

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

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

$$\Phi_\Sigma(x_1,x_2)=\frac{1}{2\pi\sqrt{\det(\Sigma)}}\iint_{y_1\leq x_1,y_2\leq x_2} e^{-\frac{1}{2}(y_1,y_2)\Sigma^{-1}(y_1,y_2)^t}dy_1dy_2.$$

Let X_n and Y_n, respectively, be random variables representing the number of encryptions and the number privileged users in a random privileged set. 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.

Theorem. With the above notation and for all the schemes considered (CST, SD and LSD), we have

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

where μ_1, μ_2 and the covariance matrix Σ are independent of n and can be computed efficiently, and $\Phi_\Sigma(x,y)$ is the distribution function of the two dimensional normal distribution with mean $(0,0)$ and covariance matrix Σ, i.e.,

$$\Phi_\Sigma(x,y) = \frac{1}{2\pi\sqrt{\det(\Sigma)}} \iint_{s\leq x, t\leq y} e^{-\frac{1}{2}(x,y)\Sigma^{-1}(x,y)^t} dsdt.$$

Lemma. For all $n \geq 0$, $|x-1| \leq 1/10$, and $|y-1| \leq 1/10$, we have

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

Lemma. For all $n \geq 0$, $|x-1| \leq 1/10$, $|y-1| \leq 1/10$, $x = e^s$ and $y = e^t$, we have,

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

where

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

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

Conclusions

• We can analyze revocation schemes for stateless receivers and provide limiting distributions for the number of encryptions and the number privileged users.

• We require a bivariate quasi-power theorem. There are now at least three problems (all coming from cryptography) where this happens. Are there more such problems? Will we need more than bivariate quasi-power theorem?

• Master theorem for nonlinear multivariate recurrences?