t-Cheater Identifiable (k, n) Threshold Secret Sharing Schemes

Kaoru KUROSAWA[1] Satoshi OBANA[1] *and* Wakaha OGATA[2]

[1] Department of Electrical and Electronic Engineering,
Faculty of Engineering, Tokyo Institute of Technology
2-12-1 O-okayama, Meguro-ku, Tokyo 152, Japan
E-mail kkurosaw@ss.titech.ac.jp
[2] Department of Computer Engineering,
Faculty of Engineering, Himeji Institute of Technology
2167 shosha, Himeji-shi, Hyougo 671–22, Japan

Abstract. In this paper, we show that there exists a t-cheater identifiable (k, n) threshold secret sharing scheme such as follows for cheating probability $\varepsilon > 0$. If $k \geq 3t + 1$, then

1. Just k participants are enough to identify who are cheaters.
2. $|V_i|$ is independent of n. That is, $|V_i| = |S|(1/\varepsilon)^{(t+2)}$, where S denotes the set of secrets and V_i denotes the set of shares of a participant P_i, respectively.

(Previously, no schemes were known which satisfy both requirements.) Further, we present a lower bound on $|V_i|$ for our model and for the model of Tompa and Woll. Our bound for the TW model is much more tight than the previous bound.

1 Introduction

In a (k, n) threshold secret sharing scheme [1, 2], a secret s is distributed by the dealer to n participants, $P_1, \ldots, P_n$, in such a way that k or more participants can recover s and $k - 1$ or less participants have no information on s. A piece of information held by P_i is called a share and it is denoted by v_i.

Various researches considered the problem of cheaters in threshold schemes. Some participants may attempt to cheat, that is, to deceive other participants by lying about shares they hold. A threshold scheme is said to be unconditionally secure against cheating if the probability of successful cheating is limited to a specified probability even if the cheaters are infinitely powerful. Assume that the dealer is honest. Then, several constructions have been given such as follows. (If the cheaters are polynomially time bounded, this problem is easily solved by using a digital signature scheme.)

McEliece and Sarwate [3] showed that Shamir's scheme itself has a cheater detection capability. Any set of $k + 2t$ participants containing at most t cheaters can detect who are cheating. They can reveal the correct secret, as well. This scheme, however, requires more than k participants to detect who are cheaters. A secret sharing scheme of [4] also requires more than k participants to detect cheaters.

T.Rabin and Ben-Or [5] showed a scheme such that each participant can always detect who are cheating with high probability. Therefore, any set of participants containing at least k honest participants can reveal the correct secret with high probability. In this scheme, however, $|V_i|$ is very large, where V_i denotes the set of shares v_i. That is, $|V_i| = |S|^{(3n-2)}$, where S denotes the set of secrets.

At the same time, Brickell and Stinson [6] showed such a nonperfect scheme in which $|V_i| = |S|^{(n+2t-3)}$. In this scheme, $k-1$ participants can have a small amount of information on the secret. In this scheme, $|V_i|$ is also an exponential function on n.

On the other hand, Tompa and Woll [7] showed a scheme such that an honest participant can detect only the fact of cheating with high probability. Honest participants, however, cannot detect who are cheating nor reveal the correct secret. Carpentieri, De Santis and Vaccaor [8] showed a lower bound on $|V_i|$ for this model.

Here, we note that cheater identifiable schemes proposed so far are either

1. $k+1$ or more participants are necessary to detect who are cheaters, or
2. $|V_i|$ is an exponential function on n.

Further, no lower bound on $|V_i|$ is known.

In this paper, we consider a model in which there are at most t cheaters. After formulation, we show that there exists a t-cheater identifiable (k, n) threshold scheme such as follows for cheating probability $\varepsilon > 0$. If $k \geq 3t+1$, then

1. Just k participants are enough to identify who are cheating.
2. $|V_i|$ is independent of n. That is, $|V_i| = |S|(1/\varepsilon)^{(t+2)}$.

The proposed scheme uses an orthogonal array $OA(t+1, n|S|, \frac{1}{\varepsilon})$. It is interesting to compare with a t error correcting BCH codes in which a generator polynomial $G(x)$ has $2t$ consecutive zeros. Further, we present a lower bound on $|V_i|$ for our model and for the model of Tompa and Woll [7]. Our bound for the TW model is much more tight than the previous bound [8].

Our scheme and bound are closely related to unconditionally secure authentication codes [12]~[16]. Especially, our bound on $|V_i|$ is derived from a bound for splitting authentication codes shown in [16].

In section 3, we give a formulation of our problem. The proposed scheme is shown in section 4. A lower bound on $|V_i|$ for our model is presented in section 5. A lower bound on $|V_i|$ for the TW model is presented in section 6.

2 Preliminaries

2.1 (k, n) threshold scheme

In a secret sharing scheme, a dealer D randomly produces $(v_1, \ldots, v_n)$ on input s, where s is a secret and v_i is called a share of the secret s. v_i is given to a participant P_i, where $i = 1, 2, \cdots, n$. Let S be the random variable induced

by s and V_i be the random variable induced by v_i, respectively. We also use S to denote the set of s and V_i to denote the set of v_i, respectively. In a (k, n) threshold scheme [1, 2], any k or more participants can recover s but no subset of less than k participants can determine any partial information on s.

Shamir's (k, n) threshold scheme [1] has error correcting capability such as follows [3].

Proposition 1. *[3] Let $i_1, \ldots, i_m$ be fixed distinct values of $GF(p)$. Let*

$$C \stackrel{\triangle}{=} \{(f(i_1), \ldots, f(i_m)) | f(x) \text{ is a degree } t \text{ polynomial over } GF(p)\}$$

where $m \geq t$. Then, C is a linear code with the minimum Hamming distance $m - t$.

Proof. It is clear that C is a linear code. Since the degree of $f(x)$ is t, the number of zeros of $f(x)$ is at most t. Therefore, the Hamming weight of $(f(i_1), \ldots, f(i_m))$ is greater than or equal to $m - t$. Further, there exists a $f(x)$ which has t zeros. Hence, the minimum Hamming weight of C is $m - t$. In a linear code, the minimum Hamming distance is equal to the minimum Hamming weight. □

[9, 10] showed that $\log_2 |V_i| \geq H(S)$ for (k, n) threshold schemes. This bound was improved by Kurosawa and Okada as follows [11]

Proposition 2. *[11] In a (k, n) threshold scheme, $|V_i| \geq |S|$ for any probability distribution on S.*

2.2 Authentication code

In the model of unconditionally secure authentication codes, there are three participants, a transmitter T, a receiver R and an opponent O. T and R share a common encoding rule e. On input a source state u, T sends a message $m = e(u)$ to R. R accepts or rejects m based on e. O tries to cheat R by an impersonation attack or a substitution attack. We assume independent probability distributions on source states and on encoding rules. In the impersonation attack, O sends m to R before T sends. O succeeds if R accepts m. This cheating probability P_I is defined by

$$P_I = \max_m \Pr[\text{R accepts } m]$$

In the substitution attack, O observes a message m transmitted by T and substitutes it with another message $\hat{m}$. This cheating probability P_S is defined by

$$P_S = \sum_m \Pr(m) \max_{\hat{m}} \Pr[\text{R accepts } \hat{m} | \text{O observed } m] \tag{2.1}$$

where the maximum is taken over $\hat{m}$ such that the source state of $\hat{m}$ is different from that of m.

Definition 3. An authentication code is called no splitting
if $|\{m \mid e(u) = m\}| = 1$ for $\forall u$ and $\forall e$. Otherwise, it is called splitting.

Let U be the set of source states and $Message$ be the set of messages respectively. Further, let

$$Message(e,u) \stackrel{\triangle}{=} \{m \mid e(u) = m\}, \qquad Message(e) \stackrel{\triangle}{=} \bigcup_u Message(e,u).$$

Proposition 4. *In an authentication code,*

(i) [14, 15] if it is no splitting,

$$P_S \geq (|U| - 1)/(|Message| - 1)$$

(ii) [16] if it is splitting,

$$P_S \geq \min_e \frac{|Message(e)| - \max_{u \in U} |Message(e,u)|}{|Message| - \min_{u \in U} |Message(e,u)|}$$

3 Formulation

In this section, we formulate our model of t-cheater identifiable (k,n) threshold scheme. In section 3∼5, we assume the following assumption.

Assumption 5. *The dealer is honest. There are at most t cheaters in n participants. (Cheaters may collude.)*

Informally, our model is defined as follows.

(T1) Completeness Any set of participants containing at least k honest participants can reveal the original secret s with high probability.
(T2) Soundness No subset of less than k participant can determine any partial information on the secret s.
(T3) Detectability There exists a Turing machine M which detects who are cheating with high probability if k or more participants open their shares.

First, we define (k,n) threshold schemes with cheaters. In what follows, let $A = \{i_1, \ldots, i_m\}$.

Definition 6. $(d_1, \ldots, d_m) \in V_{i_1} \times \cdots \times V_{i_m}$ is honest on A if

$$\Pr[V_{i_1} = d_1, \ldots, V_{i_m} = d_m] > 0$$

Definition 7. A (k,n) threshold scheme is a secret sharing scheme such as follows.
(i) If $m \geq k$, for any honest $(d_1, \ldots, d_m)$,

$$\exists s, \Pr[S = s | V_{i_1} = d_1, \ldots, V_{i_m} = d_m] = 1 \tag{3.1}$$

(ii) If $m < k$, for any honest $(d_1, \ldots, d_m)$,

$$\forall s, \Pr[S = s | V_{i_1} = d_1, \ldots, V_{i_m} = d_m] = \Pr[S = s]$$

Definition 8. If $m \geq k$ and $(d_1, \ldots, d_m)$ is honest, s is uniquely determined from eq.(3.1). Denote such s by $Secret(d_1, \ldots, d_m) = s$.

Next, we divide $V_{i_1} \times \cdots \times V_{i_n}$ into three subsets.

$$Honest(A) \stackrel{\triangle}{=} \{(d_1, \ldots, d_m) \mid (d_1, \ldots, d_m) \text{ is honest on } A\}$$

Let M be a deterministic Turing machine. For M, define

$$\begin{aligned} Dishonest_M(A) \stackrel{\triangle}{=} \{&(d_1, \ldots, d_m) \mid (d_1, \ldots, d_m) \notin Honest(A), \\ &M \text{ detects who are cheaters from } (d_1, \ldots, d_m) \text{ correctly}\} \\ Semihonest_M(A) \stackrel{\triangle}{=} \{&(d_1, \ldots, d_m) \mid (d_1, \ldots, d_m) \notin Honest(A), \\ &(d_1, \ldots, d_m) \notin Dishonest_M(A)\} \end{aligned}$$

Suppose that $P_{i_1}, \ldots, P_{i_t}$ are cheaters and they open $\hat{d_1}, \ldots, \hat{d_t}$ while the dealer distributed $(d_1, \ldots, d_m)$ to A. If $(\hat{d_1}, \ldots, \hat{d_t}, d_{t+1}, \ldots, d_m) \in Dishonest_M(A)$, M can detect who are cheaters. Successful cheating occurs if case 1 or case 2 below occurs.

(case 1) $(\hat{d_1}, \ldots, \hat{d_t}, d_{t+1}, \ldots, d_m) \in Semihonest_M(A)$. In this case, M can detect only the fact of cheating.

(case 2) $(\hat{d_1}, \ldots, \hat{d_t}, d_{t+1}, \ldots, d_m) \in Honest(A)$ and $Secret(\hat{d_1}, \ldots, \hat{d_t}, d_{t+1}, \ldots, d_m) \neq Secret(d_1, \ldots, d_m)$ (= original secret). In this case, cheaters succeed completely.

If case 1 or case 2 occurs, M cannot identify who are cheaters. This probability $Cheat_M(A)$ is formulated as follows. Let

$$\begin{aligned} &Fool_{(M,A)}(\hat{d_1}, \ldots, \hat{d_t} | d_1, \ldots, d_t) \\ &\stackrel{\triangle}{=} \{(d_{t+1}, \ldots, d_m) \mid (d_1, \ldots, d_t, d_{t+1}, \ldots, d_m) \in Honest(A), \\ &\quad \text{case 1 or 2 occurs for } (\hat{d_1}, \ldots, \hat{d_t}, d_{t+1}, \ldots, d_m)\} \end{aligned}$$

$$\begin{aligned} &Cheat_M(A | i_1, \ldots, i_t) \\ &\stackrel{\triangle}{=} \sum_{(d_1, \ldots, d_t)} \Pr[d_1, \ldots, d_t] \\ &\quad \times \max_{(\hat{d_1}, \ldots, \hat{d_t})} \sum_{Fool_{(M,A)}(\hat{d_1}, \ldots, \hat{d_t} | d_1, \ldots, d_t)} \Pr[d_{t+1}, \ldots, d_m | d_1, \ldots, d_t] \end{aligned}$$

Definition 9.

$$Cheat_M(A) \stackrel{\triangle}{=} \max\{Cheat_M(A|i_1), Cheat_M(A|i_1, i_2), \ldots, Cheat_M(A|i_1, \ldots, i_t)\}$$

Definition 10. We say that a (k, n) threshold scheme is a (t, ε) cheater identifiable (k, n) threshold scheme if there exists a deterministic Turing machine M such as follows.

$Cheat_M(A) \leq \varepsilon$ for $\forall A$ such that $|A| \geq k$.

4 Proposed scheme

In this section, we show a (t, ε) cheater identifiable (k, n) threshold scheme for $k \geq 3t + 1$ such that $|V_i|$ is independent of n. To obtain our scheme, we use an orthogonal array of strength $t + 1$ as an unconditionally secure authentication code and combine it with Shamir's (k, n) threshold scheme and the linear code of proposition 1. Each share v_i of the proposed scheme has a form of $(\alpha_i, \beta_i, \gamma_i)$. α_i is a share of the secret generated by Shamir's (k, n) threshold scheme. β_i is an authenticator for α_i of our authentication code. The key of the authentication code is encoded as a codeword $(\gamma_1, \ldots, \gamma_n)$ by the code of proposition 1. In the reconstruction phase, the key is reconstructed first. For any $i_1, \ldots, i_m$ such that $m \geq k$, $(\gamma_{i_1}, \ldots, \gamma_{i_m})$ has t-error correcting capability. Therefore, cheaters cannot forge the key. Once the key is reconstructed, a forged $(\hat{\alpha}_i, \hat{\beta}_i)$ is detected with high probability by the property of the authentication code. Thus, our scheme is (t, ε) cheater identifiable.

Definition 11. An orthogonal array $OA(t + 1, np, q)$ is a $q^{t+1} \times np$ array of q symbols such that, in any $t + 1$ columns of the array, every one of the possible q^{t+1} ordered tuples of symbols occurs in exactly one row.

In what follows, let $S = \{0, 1, \ldots, p - 1\}$, where p is a prime power.

[**Proposed Scheme**]

Dealer D produces a share $v_i = (\alpha_i, \beta_i, \gamma_i)$ such as follows for $i = 1, 2, \ldots, n$.

(C1) As in Shamir's scheme, D chooses a $(k - 1)$-th order random polynomial over $GF(p)$ such that,

$$f(x) = s + a_1 x + \cdots + a_{k-1} x^{k-1}$$

Let $\alpha_i = f(i)$ for $i = 1, 2, \ldots, n$.

(C2) Let $OA(t + 1, np, q)$ be an orthogonal array such that q is a prime power. D chooses a random number e such that $1 \leq e \leq q^{t+1}$. Let β_i be the e-th row and the
$(i - 1)p + \alpha_i$ th column element of $OA(t + 1, np, q)$.

(C3) D chooses a t-th order random polynomial over $GF(q^{t+1})$ such that

$$g(x) = e + b_1 x + \cdots + b_t x^t$$

Let $\gamma_i = g(i)$ for $i = 1, 2, \ldots, n$.

Remark. We assume that $OA(t + 1, np, q)$ is publicly known.

Theorem 12. *The above scheme is a (t, ε) cheater identifiable (k, n) threshold scheme if $k \geq 3t + 1$, where $\varepsilon = 1/q$.*

Proof. (1) Def.7 is satisfied because α_i is a share of Shamir's (k, n) threshold scheme.
(2) Suppose that $A = \{i_1, \ldots, i_m\}$ and $m \geq k$. Let $d_j = (\alpha_{i_j}, \beta_{i_j}, \gamma_{i_j})$ for $j = 1, 2, \ldots, m$.
Clearly,

$$\mathit{Honest}(A) = \{(d_1, \ldots, d_m) \mid (d_1, \ldots, d_m) \text{ satisfies (C1)} \sim \text{(C3).}\}$$

We show that there exists a deterministic Turing machine M such that

$$\mathit{Dishonest}_M(A) = \{(d_1, \ldots, d_m) \mid (d_1, \ldots, d_m) \text{ doesn't satisfy (C2) or (C3).}\} \tag{4.1}$$

Suppose that $(d_1, \ldots, d_m)$ doesn't satisfy (C3). That is,

$$(\gamma_{i1}, \ldots, \gamma_{im}) \neq (g(i_1), \ldots, g(i_m)),$$

where $g(x)$ is a degree t polynomial chosen by D. From Proposition 1, $\{(g(i_1), \ldots, g(i_m))\}$ is a linear code with the Hamming distance $d = m - t$. In our case,

$$m \geq k \geq 3t + 1$$

Hence

$$d \geq 3t + 1 - t = 2t + 1$$

Therefore, there is a deterministic algorithm which can identify t errors in $(\gamma_{i1}, \ldots, \gamma_{im})$.

Also, there is a deterministic algorithm which can recover $g(x)$. Then, e is reconstructed. Now, we see that there is a deterministic algorithm which finds e in any case. Once e is found, it is easy to detect which β_i violates (C2).

Thus, there exists a deterministic Turing machine which detects who are cheaters if $(d_1, \ldots, d_m)$ doesn't satisfy (C2) or (C3). Let M be a deterministic Turing machine which detects who are cheaters if $(d_1, \ldots, d_m)$ doesn't satisfy (C2) or (C3). Then, eq.(4.1) holds.
(3) Finally, we prove that $\mathit{Cheat}_M(A) \leq 1/q$ for the above M. Suppose that $P_{i_1}, \ldots, P_{i_t}$ are cheaters and they open $\hat{d_1}, \ldots, \hat{d_t}$ while the dealer distributed $(d_1, \ldots, d_m)$ to A. Let

$$\begin{aligned} &F_{(M,A)}(\hat{d_1}, \ldots, \hat{d_t} | d_1, \ldots, d_t) \\ &\quad \triangleq \{(d_{t+1}, \ldots, d_m) \mid (d_1, \ldots, d_m) \in \mathit{Honest}(A), \\ &\qquad\qquad (\hat{d_1}, \ldots, \hat{d_t}, d_{t+1}, \ldots, d_m) \notin \mathit{Dishonest}_M(A)\} \end{aligned}$$

Then, it is easy to see that

$$F_{(M,A)}(\hat{d_1}, \ldots, \hat{d_t} | d_1, \ldots, d_t) \supseteq \mathit{Fool}_{(M,A)}(\hat{d_1}, \ldots, \hat{d_t} | d_1, \ldots, d_t)$$

Let

$$\begin{aligned} F1_A(d_1, \ldots, d_t) &\triangleq \{(d_{t+1}, \ldots, d_m) | (d_1, \ldots, d_m) \in \mathit{Honest}(A)\} \\ F2_A(\hat{d_1}, \ldots, \hat{d_t}) &\triangleq \{(d_{t+1}, \ldots, d_m) | (\hat{d_1}, \ldots, \hat{d_t}, d_{t+1}, \ldots, d_m) \notin \mathit{Dishonest}_M(A)\} \end{aligned}$$

Then

$$F_{(M,A)}(\hat{d}_1,\ldots,\hat{d}_t|d_1,\ldots,d_t) = F1_A(d_1,\ldots,d_t) \cap F2_A(\hat{d}_1,\ldots,\hat{d}_t)$$

Let's compute $|F1_A(d_1,\ldots,d_t)|$. Fix an honest $d_1,\ldots,d_t$ arbitrarily. From the definition of $OA(t+1,np,q)$ and since $\deg g(x) = t$, there are q rows which matches $(d_1,\ldots,d_t)$. That is,

$$|\{e \mid \beta_{i_j} = \text{ the } (e,(i_j-1)p+\alpha_{i_j}) \text{ element of } OA(t+1,np,q), 1 \le j \le t\}| = q$$

For each e, $g(x)$ is uniquely determined. Then, γ_{i_j} is uniquely determined for $t+1 \le j \le m$. On the other hand, there are p^{k-t} possible $f(x)$ which matches $(d_1,\ldots,d_t)$. For each $f(x)$ and e, $(\alpha_{i_j},\beta_{i_j})$ is uniquely determined for $t+1 \le j \le m$. Therefore, $|F1_A(d_1,\ldots,d_t)| = qp^{k-t}$.
Similarly, we see that

$$|F_{(M,A)}(\hat{d}_1,\ldots,\hat{d}_t|d_1,\ldots,d_t)| = 0 \text{ or } p^{k-t}$$

Then,

$$\begin{aligned}
&\max_{(\hat{d}_1,\ldots,\hat{d}_t)} \sum_{Fool_{(M,A)}(\hat{d}_1,\ldots,\hat{d}_t|d_1,\ldots,d_t)} \Pr[d_{t+1},\ldots,d_m|d_1,\ldots,d_t] \\
&\le \max_{(\hat{d}_1,\ldots,\hat{d}_t)} \sum_{F_{(M,A)}(\hat{d}_1,\ldots,\hat{d}_t|d_1,\ldots,d_t)} \Pr[d_{t+1},\ldots,d_m|d_1,\ldots,d_t] \\
&= \max_{(\hat{d}_1,\ldots,\hat{d}_t)} \frac{|F_{(M,A)}(\hat{d}_1,\ldots,\hat{d}_t|d_1,\ldots,d_t)|}{|F1_A(d_1,\ldots,d_t)|} \\
&= \frac{p^{k-t}}{qp^{k-t}} = \frac{1}{q}
\end{aligned}$$

Therefore,

$$Cheat_M(A|i_1,\ldots,i_t) \le \sum_{(d_1,\ldots,d_t)} \Pr[d_1,\ldots,d_t] \times (1/q) = 1/q$$

Similarly,

$$Cheat_M(A|i_1,\ldots,i_j) \le 1/q, \quad 1 \le j \le t-1$$

Therefore,

$$Cheat_M(A) \le 1/q$$

□

In this scheme,

$$\begin{aligned}
\log_2|V_i| &= \log_2|\alpha_i| + \log_2|\beta_i| + \log_2|\gamma_i| \\
&= \log_2|p| + \log_2|q| + \log_2|q^{t+1}| \\
&= \log_2|S| + (t+2)\log_2(1/\varepsilon)
\end{aligned}$$

Equivalently,

$$|V_i| = |S|(1/\varepsilon)^{t+2}$$

Thus, $|V_i|$ is independent of n.

5 Lower bound on $|V_i|$

From proposition 2, $|V_i| \geq |S|$ in any (k, n) threshold scheme. In general, a perfect secret sharing scheme is called ideal if $|V_i| = |S|$ for $\forall i$. First, we show a refinement of this bound and the concept. Let $A = \{i_1, \ldots, i_k\}$ ($|A| = k$). Let d_j be the share of P_{i_j}, where $j = 1, \ldots, k$.

Theorem 13. *In a (k, n) threshold scheme,*

$$|\{d_1 \mid \Pr(s, d_1, d_2, \ldots, d_k) > 0\}| \geq 1$$

for any secret s and any honest $(d_2, \ldots, d_k)$.

Proof. From Def.7 (ii), for any honest $(d_2, \ldots, d_k)$, s can take any value of S. From Def.7 (i), each s of S must be determined by some d_1 together with this $(d_2, \ldots, d_k)$. This means that Theorem 13 holds. □

Definition 14. A (k, n) threshold scheme is c-compact on (A, i_1) if

$$|\{d_1 \mid \Pr(s, d_1, d_2, \ldots, d_k) > 0\}| = c$$

for any secret s and any honest $(d_2, \ldots, d_k)$.

Next, for simplicity, suppose that only P_{i_1} is a cheater. Let $Cheat(A|i_1)$ be the cheating probability that the case 2 occurs for $(\hat{d_1}, d_2, \ldots, d_k)$ for a forged $\hat{d_1}$. Formally,

$$\begin{aligned} Fool_A(\hat{d_1}|d_1) &\stackrel{\triangle}{=} \{(d_2, \ldots, d_k) \mid (d_1, d_2, \ldots, d_k) \in Honest(A), \\ &\qquad \text{case 2 occurs for } (\hat{d_1}, d_2, \ldots, d_k)\} \\ Cheat(A|i_1) &\stackrel{\triangle}{=} \sum_{d_1} \Pr[d_1] \times \max_{\hat{d_1}} \sum_{Fool_A(\hat{d_1}|d_1)} \Pr[d_2, \ldots, d_k|d_1] \end{aligned}$$

Note that the subscript M is dropped in the above definitions because case 2 is independent of M.

Theorem 15. *Suppose that a (k, n) threshold scheme is c-compact on (A, i_1) for some $A \ni i_1$. If $Cheat(A|i_1) \leq \varepsilon$,*

$$|V_{i_1}| \geq c\left(\frac{|S| - 1}{\varepsilon} + 1\right) \tag{5.1}$$

Proof. Consider a splitting authentication code such as follows (see subsection 2.2). The receiver R has an encoding rule $e = (d_2, \ldots, d_k)$ such that $(d_2, \ldots, d_k)$ is honest. R accepts a message $m = d_1$ if $(d_1, d_2, \ldots, d_k)$ is honest. The source state u conveyed by d_1 is a secret s such that $s = Secret(d_1, d_2, \ldots, d_k)$. (see Def.8.) P_{i_1} is the opponent. He observes d_1 and substitutes d_1 with another $\hat{d_1}$ (substitution attack). Then

$$Cheat(A|i_1) = \sum \Pr[d_1] \max_{\hat{d}_1} \sum_{Fool_A(\hat{d}_1|d_1)} \Pr[d_2, \ldots, d_k|d_1]$$
$$= \sum \Pr[d_1] \max_{\hat{d}_1} \Pr[\text{R accepts } \hat{d}_1 | P_{i_1} \textit{ observed } d_1] \qquad (5.2)$$

where the maximum is taken over $\hat{d}_1$ such that the source state (secret) determined by $\hat{d}_1$ is different from that of d_1. Now, we see that $Cheat(A|i_1)$ is equal to the substitution attack probability P_S (compare eq.(5.2) with eq.(2.1)). In this authentication code,

$$|Message| = |\{d_1\}| = |V_{i_1}|$$
$$|Message((d_2, \ldots, d_k), s)| = |\{d_1 \mid \Pr(s, d_1, d_2, \ldots, d_k) > 0\}| = c$$
$$|Message((d_2, \ldots, d_k))| = |\bigcup_s Message((d_2, \ldots, d_k), s)| = c|S|$$

Then from proposition 4 (ii),

$$\varepsilon \geq Cheat(A|i_1) = P_S \geq \min_{(d_2,\ldots,d_k)} \frac{c|S| - c}{|V_{i_1}| - c} = \frac{c(|S| - 1)}{|V_{i_1}| - c}$$

Therefore, eq.(5.1) is obtained □

Finally, we show a lower bound on $|V_i|$ of our model.

Corollary 16. *If a (t, ε) cheater identifiable (k, n) threshold scheme is c-compact on (A, i_1) for some $A \ni i_1$,*

$$|V_{i_1}| \geq c\left(\frac{|S| - 1}{\varepsilon} + 1\right) \qquad (5.3)$$

Proof. There exists a deterministic Turing machine M such that $Cheat_M(A|i_1) \leq Cheat_M(A) \leq \varepsilon$. Clearly, $Cheat(A|i_1) \leq Cheat_M(A|i_1)$. Therefore, $Cheat(A|i_1) \leq \varepsilon$. Then, from Theorem 15, we obtain this corollary. □

6 Lower bound for TW model

Tompa and Woll [7] showed a (k, n) threshold scheme such that an honest participant can detect only the fact of cheating with high probability. Honest participants, however, cannot detect who are cheating nor reveal the correct secret. Carpentieri, De Santis and Vaccaor [8] showed a lower bound on $|V_i|$ for this model.

In this section, we show a much more tight lower bound on $|V_i|$ for the model of Tompa and Woll. (We don't assume Assumption 5 in this section while we use the same notation as before.)

In the model of Tompa and Woll [7], the cheating probability, P_{TW}, is defined as the probability that from $k - 1$ forged shares $d'_1, \ldots, d'_{k-1}$ and any d_k, the

secret s' reconstructed is legal but not a correct one. Formally, it should be defined as follows.

$$\begin{aligned} \mathit{cheated} &\stackrel{\triangle}{=} P_k \text{ has } d_k \text{ such that } (d'_1, \ldots, d'_{k-1}, d_k) \text{ is honest and} \\ &\qquad \mathit{Secret}(d_1, \ldots, d_k) \neq \mathit{Secret}(d'_1, \ldots, d'_{k-1}, d_k). \\ P_{TW} &\stackrel{\triangle}{=} \sum_{d_1,\ldots,d_{k-1}} \Pr(d_1, \ldots, d_{k-1}) \max_{d'_1,\ldots,d'_{k-1}} \Pr(\mathit{cheated} | d_1, \ldots, d_{k-1}) \end{aligned}$$

For a technical reason, [8] defined the following probability.

$$P_{over} \stackrel{\triangle}{=} \sum_{d_1,\ldots,d_{k-1},s} \Pr(d_1, \ldots, d_{k-1}, s) \max_{d'_1,\ldots,d'_{k-1}} \Pr(\mathit{cheated} | d_1, \ldots, d_{k-1}, s)$$

Lemma 17. *For any function* $f(x_1, x_2)$,

$$\sum_{x_2} \max_{x_1} f(x_1, x_2) \geq \max_{x_1} \sum_{x_2} f(x_1, x_2)$$

Proof. For $\forall \hat{x_1}$,

$$\sum_{x_2} \max_{x_1} f(x_1, x_2) \geq \sum_{x_2} f(\hat{x_1}, x_2)$$

□

Lemma 18. $P_{over} \geq P_{TW}$

Proof. From Def.7 (ii), $\Pr(d_1, \ldots, d_{k-1}, s) = \Pr(d_1, \ldots, d_{k-1}) \Pr(s)$. Then,

$$\begin{aligned} P_{over} &= \sum_{d_1,\ldots,d_{k-1},s} \Pr(d_1, \ldots, d_{k-1}) \Pr(s) \max_{d'_1,\ldots,d'_{k-1}} \Pr(\mathit{cheated} | d_1, \ldots, d_{k-1}, s) \\ &= \sum_{d_1,\ldots,d_{k-1},s} \Pr(d_1, \ldots, d_{k-1}) \\ &\quad \times \max_{d'_1,\ldots,d'_{k-1}} \Pr(S = s \text{ and } \mathit{cheated} | d_1, \ldots, d_{k-1}) \\ &= \sum_{d_1,\ldots,d_{k-1}} \Pr(d_1, \ldots, d_{k-1}) \\ &\quad \times \sum_s \max_{d'_1,\ldots,d'_{k-1}} \Pr(S = s \text{ and } \mathit{cheated} | d_1, \ldots, d_{k-1}) \end{aligned}$$

From lemma 17,

$$\begin{aligned} P_{over} &\geq \sum_{d_1,\ldots,d_{k-1}} \Pr(d_1, \ldots, d_{k-1}) \\ &\qquad \times \max_{d'_1,\ldots,d'_{k-1}} \sum_s \Pr(S = s \text{ and } \mathit{cheated} | d_1, \ldots, d_{k-1}) \\ &= P_{TW} \end{aligned}$$

□

[7] showed a scheme such that $P_{over} \leq \varepsilon$ and

$$|V_i| \geq \{(|S|-1)(k-1)/\varepsilon + k\}^2$$

[8] showed that, if $P_{over} \leq \varepsilon$ and S is uniformly distributed, then

$$|V_i| \geq |S|/\varepsilon$$

From lemma 18, these results [7, 8] can be restated as follows.

1. There exists a (k, n) threshold scheme such that $P_{TW} \leq \varepsilon$ and
$$|V_i| \geq \{(|S|-1)(k-1)/\varepsilon + k\}^2 \tag{6.1}$$
2. If $P_{TW} \geq \varepsilon$ and S is uniformly distributed, then
$$|V_i| \geq |S|/\varepsilon \tag{6.2}$$

However, there is a big gap between eq.(6.1) and eq.(6.2). In what follows, we show a much more tight lower bound on $|V_i|$ than eq.(6.2).

Lemma 19. $P_{TW} \geq \mathit{Cheat}(A|i_1)$.

Proof.

$$\begin{aligned}
&P_{TW} \\
&= \sum_{d_1} \Pr(d_1) \sum_{d_2,\ldots,d_{k-1}} \Pr(d_2,\ldots,d_{k-1}|d_1) \max_{d'_1,\ldots,d'_{k-1}} \Pr(\mathit{cheated}|d_1,\ldots,d_{k-1}) \\
&= \sum_{d_1} \Pr(d_1) \sum_{d_2,\ldots,d_{k-1}} \max_{d'_1,\ldots,d'_{k-1}} \Pr(\mathit{cheated} \text{ and } P_2,\ldots,P_k \text{ has } d_2,\ldots,d_{k-1}|d_1)
\end{aligned}$$

Let

$$\begin{aligned}
B &\stackrel{\triangle}{=} P_2,\ldots,P_{k-1} \text{ has } d_2,\ldots,d_{k-1}. \\
C &\stackrel{\triangle}{=} P_k \text{ has } d_k \text{ such that } (d'_1 d_2,\ldots,d_k) \text{ is honest} \\
&\quad \text{and } \mathit{Secret}(d_1,\ldots,d_k) \neq \mathit{Secret}(d'_1,d_2,\ldots,d_k).
\end{aligned}$$

Then,

$$\max_{d'_1,\ldots,d'_{k-1}} \Pr(\mathit{cheated} \text{ and } B|d_1) \geq \max_{d'_1} \Pr(C \text{ and } B|d_1)$$

Therefore, from lemma 17,

$$\begin{aligned}
P_{TW} &\geq \sum_{d_1} \Pr(d_1) \sum_{d_2,\ldots,d_{k-1}} \max_{d'_1} \Pr(C \text{ and } B|d_1) \\
&\geq \sum_{d_1} \Pr(d_1) \max_{d'_1} \sum_{d_2,\ldots,d_{k-1}} \Pr(C \text{ and } B|d_1) \\
&= \sum_{d_1} \Pr(d_1) \max_{d'_1} \sum_{Fool_A(\hat{d}_1|d_1)} \Pr[d_2,\ldots,d_k|d_1] \\
&= \mathit{Cheat}(A|i_1)
\end{aligned}$$

□

Corollary 20. *Suppose that a (k,n) threshold scheme is c-compact on (A,i_1) for some $A \ni i_1$. If $P_{TW} \leq \varepsilon$, then*

$$|V_{i_1}| \geq c\left(\frac{|S|-1}{\varepsilon}+1\right) \tag{6.3}$$

Proof. From lemma 19, if $P_{TW} \leq \varepsilon$, then $Cheat(A|i_1) \leq \varepsilon$. Then, from Theorem 15, we obtain this corollary. □

Eq.(6.3) is more tight than eq.(6.2) if $c \geq 2$. The (k,n) threshold scheme of Tompa and Woll [7] is c-compact on $\forall(A,i)$ such that $c \geq (|S|-1)/(k-1)/\varepsilon$. Then, our bound becomes

$$|V_i| \geq \{(|S|-1)/\varepsilon+1\}(|S|-1)(k-1)/\varepsilon$$

This bound is much closer to eq.(6.1) than eq.(6.2).

Remark. In Theorem 15, Corollary 16 and Corollary 20, we can eliminate "c-compact" to obtain a more general bound on $|V_i|$, which is as general as Proposition 4 (ii). The details will be given in the final paper.

References

1. A.Shamir, How to share a secret, Comm.ACM, 22(1979), pp.612–613.
2. G.R.Blakley, Safeguarding cryptographic keys, Proc. National Computer Conference, AFIPS Conference Proceedings, 48(1979), pp.313–317.
3. R.J.McEliece and D.V.Sarwate, On sharing secrets and Reed-Solomon codes, Comm.ACM, 24(1981), pp.583–584.
4. G.Simmons, Robust shared secret schemes or "how to be sure you have the right answer even though you don't know the question," Congr.Numer., 68(1989), pp.215–248.
5. T.Rabin and M.Ben-Or, Verifiable secret sharing and multiparty protocols with honest majority, Proc. 21st ACM Symposium on Theory of Computing (1989), pp.73–85.
6. E.F.Brickell and D.R.Stinson, The Detection of Cheaters in Threshold Schemes, SIAM J. DISC. MATH, Vol.4, No.4, Nov.1991, pp.502–510.
7. M.Tompa and H.Woll, How to share a secret with cheaters, Journal of Cryptology, vol.1(1988), pp.133–138.
8. Marco Carpentieri, Alfredo De Santis and Ugo Vaccaro, Size of Shares and Probability of Cheating in Threshold Schemes, Proceedings of Eurocrypt'93, Lecture Notes in Computer Science, LNCS 765, Springer Verlag (1993), pp.118–125.
9. E.D.Karnin, J.W.Greene, and M.E.Hellman, On Secret Sharing Systems, IEEE Trans. on Inform. Theory, Vol.IT-29 (1983), pp.35–41.
10. R.M.Capocelli, A.De Santis, L.Gargano and U.Vaccaro, On the size of shares for secret sharing schemes, Proceedings of Crypto'91, Lecture Notes in Computer Science, LNCS 576, Springer Verlag (1991), pp.101-113.
11. K.Kurosawa and K.Okada, Combinatorial interpretation of secret sharing schemes, In Pre-Proceedings of Asiacrypt'94 (1994), pp.38–48.

12. G.J.Simmons, A survey of Information Authentication, in Contemporary Cryptology, The science of information integrity, ed. G.J.Simmons, IEEE Press, New York (1992).
13. G.J.Simmons, Message authentication: a game on hypergraphs, Congr. Numer. 45 (1984), pp.161–192.
14. D.R.Stinson, Some constructions and bounds for authentication codes, Journal of Cryptology, vol.1 (1988), pp.37–51.
15. J.L.Massey, Cryptography – a selective survey, in Digital Communications, North-Holland (pub.) (1986), pp.3–21.
16. M.De Soete, New Bounds and Constructions for Authentication/Secrecy Codes with Splitting, Journal of Cryptology, vol.3, no.3 (1991), pp.173–186.

This article was processed using the LaTeX macro package with LLNCS style