微小な復号誤り確率を許容する効率的な秘密分散法の研究（継続）

古　賀　弘　樹　　筑波大学大学院システム情報工学研究科准教授

1 はじめに

秘密分散法はインターネット社会の安全を高める技術である。秘密分散法では、秘密情報を**シェア**と呼ばれる分散情報に変換して保管する。ある特定の分散情報が集まれば秘密情報が復元できるが、それ以外は秘密情報が一切漏れない。特に、しきい値法と呼ばれる秘密分散法では、秘密情報が復元できるかどうかが、集まった分散情報の数で決定する。秘密分散法は Shamir[1]と Blakley[2]によって独立に提案された後、様々な形に拡張されて発展してきた。中でも、McEliece と Sarwate[3]は、シェアを改ざんする不正者の存在のもとでの秘密分散法を考察した。その後、文献[4][5][6]などによってこの問題が研究されている。ところが、従来手法のほとんどが Shamir のしきい値法を用いており、不正が成功する確率を厳密に評価することは難しく、特に不正の成功確率に関する上界の議論は全くなかった。

本研究では、情報理論的な立場から不正者が存在する秘密分散法の問題を扱い、その基礎的性質を明らかにし、不正者の攻撃に対して耐性をもつ新しい秘密分散の方式を構築することを目的としている。

2 不正者が存在する(t,m)しきい値法

秘密情報 S を有限集合 $\mathcal{S}$ に値をとる確率変数であるとし、$\mathcal{P}=\{1,2,\ldots,m\}$を参加者の集合とする。秘密分散法では、ディーラが、秘密情報 S を m 個のシェア $X_1, X_2, \ldots, X_m$に変換する。特に、(t,m)しきい値法と呼ばれる秘密分散法では、秘密情報 S は任意の t 個のシェアから復号でき、かつ、t-1 個以下のどのシェアからも秘密情報が漏れないという性質をもっている。すなわち t は、複数個のシェアが集まったときに、秘密情報が復元できるか、できないか、を定める「しきい値」としての意味をもつ。 Shamir 法[1]は一般の $2 \leqq t \leqq m$ に対して(t,m)しきい値法を実現することができ、t=m の特別の場合には、Karnin らの方法[7]ても(m,m)しきい値法を実現することができる。どちらの方法もシェアの生成には一様乱数を用いる。

本研究では、不正者が存在する状況での(t,m)しきい値法を考える。我々は、不正者の攻撃として次の 2 通りを考える。

(A) 不正者は参加者集合 $\mathcal{P}$ には属さず、参加者 i($1 \leqq i \leqq m$)になりすまして、他の t-1 人の参加者のもつシェアと、自分で偽造したシェア X_i' をもとに、秘密情報 S' を復元し、他の t-1 人の参加者に自分が参加者 i であると認めさせることを目的とする。

(B) 不正者は参加者集合 $\mathcal{P}$ に属する参加者 i($1 \leqq i \leqq m$)であり、参加者 i は自分のもつシェア X_i をもとに不正なシェア X_i'を偽造する。参加者 i は、他の t-1 人の参加者のもつシェアと、偽造したシェア X_i' をもとに、もとの秘密情報とは異なる秘密情報 S' を復元し、他の t-1 人の参加者を騙すことを目的とする。

本稿では、(A)を**なりすまし攻撃**(impersonation attack)、(B)を**改ざん攻撃**(substitution attack)と呼ぶ。(A)のなりすまし攻撃の場合は復号時に $\mathcal{S}$ の元が出力されれば成功とみなされる。不正者は m 個のシェア $X_1, X_2, \ldots, X_m$ と独立にシェア X_i'を偽造することになる。(B)の改ざん攻撃の場合は、$S' \neq S$ かつ $S' \in \mathcal{S}$ のときに成功とみなされる。シェアの改ざんを行う参加者 i は自分のシェアを見ることができるので、一般に X_j ($j \neq i$), X_i, X_i'はこの順に Markov 連鎖をなす。

以下では、３つの異なる問題設定のもとで、これらの攻撃に対して安全な(t,m)しきい値法を構成し、それらの基本的な性質を明らかにする。

3 なりすまし攻撃に対して安全な無記憶情報源に対する(2,2)しきい値法定理

本節では、秘密情報が無記憶情報源から出力される長さ n の系列 $S^n=S_1S_2\ldots S_n \in \mathcal{S}^n$ であり、簡単のために(t,m)=(2,2)の場合を考える。符号器は、一様乱数 $E_n \in \mathcal{E}_n$ を用いて秘密情報 S^n を２つのシェア X_n, Y_n に変換する。X_n, Y_n はそれぞれ有限集合 $\mathcal{X}_n$, $\mathcal{Y}_n$ に値をとるとし、X_n は参加者１に、Y_n は参加者２にそれぞれ配布される。他方、復号器は、秘密情報 S^n を２つのシェア X_n と Y_n から１に近い確率で復号する。我々は、復号誤り

確率 P_e が n→∞で 0 になることを要請する。また、(2,2)しきい値法としての要請を満たすため、n が十分大きいときには S^n の情報は X_n, Y_n のどちらか一方からはほとんど漏れないという要請も課す。また、不正者のなりすまし攻撃だけを考え、参加者１へのなりすましが成功する確率を $P_{I,1}$, 参加者２へのなりすましが成功する確率を $P_{I,2}$ と書く。$P_{I,1}$, $P_{I,2}$ は小さければ小さいほどよい。

これらの条件を満たす符号器と復号器が存在するために必要はシェア X_n, Y_n のレートおよび一様乱数 En のレートを求めることを考える。このために、２つのシェアの**相関レベル**を $I(X_n;Y_n)/n$ の極限値として定義する。ここに $I(X_n;Y_n)$ は X_n と Y_n の相互情報量である。

我々は次の定理を得た。

定理１（逆定理[8]）２つのシェアの相関レベルが r に等しく、上記の条件をすべて満たす任意の符号器と復号器に対して、以下の式が成り立つ。

$$\liminf_{n\to\infty} \frac{1}{n}\log |X_n| \geqq H(S) + r \tag{1}$$

$$\liminf_{n\to\infty} \frac{1}{n}\log |Y_n| \geqq H(S) + r \tag{2}$$

$$\liminf_{n\to\infty} \frac{1}{n}\log |\epsilon_n| \geqq H(S) + r \tag{3}$$

$$\limsup_{n\to\infty} \max\{-\frac{1}{n}\log P_{I,1}, -\frac{1}{n}\log P_{I,2}\} \leqq r \tag{4}$$

ここに H(S)は情報源のエントロピーである。

定理１の式(1),(2)はシェアのレートが漸近的には H(S)+r 以下にはできないこと、式(3)は一様乱数のレートが漸近的には H(S)+r 以下にはできないことを示している。式(4)は、なりすまし攻撃が、n が十分大きいときには 2^{-nr} 以上の確率で成功することを示している。式(4)の性質を示すときには、情報理論的な仮説検定で用いられる不等式を利用する。

次の定理は、定理１で得られた限界が漸近的に達成可能であることを示している。

定理２（順定理[8]）r を任意の非負定数とすると、次の４つの式を満たす、相関レベル r の符号器と復号器が構成できる。

$$\limsup_{n\to\infty} \frac{1}{n}\log |X_n| \leqq H(S) + r \tag{5}$$

$$\limsup_{n\to\infty} \frac{1}{n}\log |Y_n| \leqq H(S) + r \tag{6}$$

$$\limsup_{n\to\infty} \frac{1}{n}\log |\epsilon_n| \leqq H(S) + r \tag{7}$$

$$\liminf_{n\to\infty} \min\{-\frac{1}{n}\log P_{I,1}, -\frac{1}{n}\log P_{I,2}\} \geqq r \tag{8}$$

定理２の証明における符号器と復号器の構成は簡単である。まず情報源出力 S^n をエントロピー程度のレートで記述できるように、固定長符号化で符号化する。次に、その符号語を、Karnin らの方式[7]で(2,2)しきい値法に変換した後、２つのシェアに長さ nr の一様乱数を連接するという方式である。２つのシェアの復号は、一様乱数が連接された部分が同一であれば２つのシェアを受理し、異なれば２つのシェアが不正であるとして棄却する。不正者が参加者へのなりすましに成功するのはこの一様乱数部分を正確に偽造できたときに限られ、この確率は 2^{-nr} となる。

4 なりすまし攻撃に対して安全な一般情報源に対する(2,2)しきい値法

第３節で考えた(2,2)しきい値法の枠組みでは、秘密情報 S^n は無記憶情報源の出力であることを仮定していた。本節では、この「無記憶情報源」という仮定を外すことを考える。実は、定理１と定理２の内容は、秘密情報 S^n が定常エルゴード情報源から出力される場合に拡張できることは、定理の証明から比較的容易に

分かる[8]。この拡張では、式(1)—(3)、式(5)—(7)のエントロピーは情報源のエントロピーレートで置き換わる。

本節では、一般情報源から出力される場合を考える。一般情報源は定常エルゴード情報源を含む極めて広範な情報源クラスである。一般情報源に対しては、これまで情報源符号化や通信路符号化、仮説検定等の情報理論の問題に対して、従来とは異なる形の符号化定理が導出されており[9]、また、研究代表者によって、(t,m)しきい値法に対する新たな符号化定理も導出されている[10]。一般情報源を扱う場合の特徴は、従来型の情報理論でしばしば用いられるエントロピーや相互情報量などの概念が操作的な意味をもたなくなることであり、この意味において、第３節で扱った(2,2)しきい値法の問題の新たな面白さを見いだすことが可能になる。

我々は、確率変数の列{Zn}に対して、

$$p\liminf_{n\to\infty} Z_n = \sup\{\beta;\lim_{n\to\infty}\Pr(Z_n \ge \beta) = 1\}$$

(9)

$$p\liminf^{*}_{n\to\infty} Z_n = \sup\{\beta;\limsup_{n\to\infty}\Pr(Z_n \ge \beta) = 1\} \tag{10}$$

と定義する。式(9)は文献[9]等で定義される確率的下極限である。さらに、

$$\underline{I}(X,Y) = p\liminf_{n\to\infty}\frac{1}{n}\log\frac{P_{X_n}(X_n)P_{Y_n}(Y_n)}{P_{X_nY_n}(X_n,Y_n)} \tag{11}$$

$$\underline{I}^{*}(X,Y) = p\liminf^{*}_{n\to\infty}\frac{1}{n}\log\frac{P_{X_n}(X_n)P_{Y_n}(Y_n)}{P_{X_nY_n}(X_n,Y_n)} \tag{12}$$

と定める。ここに $P_{Xn}(X_n)$ は X_n が生成される確率を表し $P_{Yn}(Y_n)$, $P_{XnYn}(X_n, Y_n)$ も同様である。これらの記法のもとで、定理１の拡張にあたる次の定理を得た。

定理３（[11]） 非負整数列{r_n}に対して

$$p\liminf_{n\to\infty}\frac{1}{n}\log\frac{P_{X_n}(X_n)P_{Y_n}(Y_n)}{2^{nr}P_{X_nY_n}(X_n,Y_n)} \ge 0$$

を満たす任意の符号器と復号器に対して、以下の４つの式が成り立つ。

$$p\liminf_{n\to\infty}\frac{1}{n}\log\frac{P_{S^n}(S^n)}{2^{nr}P_{X_n}(X_n)} \ge \underline{I}(X;Y) \tag{13}$$

$$p\liminf_{n\to\infty}\frac{1}{n}\log\frac{P_{S^n}(S^n)}{2^{nr}P_{Y_n}(Y_n)} \ge \underline{I}(X;Y) \tag{14}$$

$$p\liminf_{n\to\infty}\frac{1}{n}\log(|\varepsilon_n| P_{S^n}(S^n)) \ge \underline{I}(X;Y) \tag{15}$$

$$p\limsup_{n\to\infty}\max\{-\frac{1}{n}\log P_{I,1}, -\frac{1}{n}\log P_{I,2}\} \ge \underline{I}^{*}(X;Y) \tag{16}$$

情報源が無記憶であり、前節の意味で２つのシェアの相関レベルが r であれば、式(13)—(16)の右辺はいずれも r に一致する。ところが、より一般的な場合では、式(13)—(15)と式(16)の右辺に現われる量は異なる。*印がついた量は楽観的符号化[11]と関連する量であり、楽観的符号化以外の文脈では従来ほとんど出てきていない量であるが、今回新たに秘密分散法と楽観的符号化の関連が明らかになった。

定理２に相当する拡張も導出することができるが、今回は割愛する。詳細は文献[11]を見られたい。

5 改ざん攻撃に対して安全な(m,m)しきい値法

本節では、不正者の改ざん攻撃に対して耐性をもつ秘密分散法を構成する。最初に、符号器と復号器は鍵 K を共有できることを仮定する。鍵 K は有限集合 $\mathcal{K}$ 上の一様分布に従うとする。秘密情報を S とし、符号器だけが使うことができる乱数を E と書く。S と E は相関をもってよいが、組(S,E)と K は独立であるとする。簡単のため、シェアの数が２の場合のシェア生成アルゴリズムを述べる。以下では、GF(p)={0,1,…, p-1}で p 個の元をもつ有限体を表し、$\mathcal{S}$={1,2,…,q-1}を GF(p)の部分集合であるとする。秘密情報 S と乱数 E は集合$\mathcal{S}$に値をとるとし、２つのシェア X, Y は GF(p) -{0}に値をとるとする。鍵 K は、(GF(p)-{0})上の一様乱数 U,V を用いて K=(U,V)と定まるものとする。すべての演算は GF(p)で定義されているものを用いる。

シェア生成アルゴリズム

入力：秘密情報 S, 鍵 K=(U,V)　　出力：シェア X, Y

(E1) E を $\mathcal{S}$ - {S}から一様ランダムに選ぶ。

(E2) X = U^{-1} E, Y = V^{-1} (S-E)を出力する (U^{-1},V^{-1}は、それぞれ U,V の乗算に関する逆元)。

上のシェア生成アルゴリズムにおいて、U^{-1}, V^{-1}は必ず存在し、E に関する仮定から X, Y も(GF(p)-{0})に値をとり、UX +VY = S が必ず成り立つことに注意する。

復号アルゴリズム

入力：２つのシェア X',Y', 鍵 K=(U,V)　　出力：$\mathcal{S}$の元または reject

(D1) UX' + VY' を計算する。

(D2) UX' + VY' ∈ $\mathcal{S}$ならばその値を S として出力する。そうでなければ reject を出力する。

シェアに改ざんがないとき、すなわち(X',Y')=(X,Y)のときは、上記の復号アルゴリズムでは秘密情報 S が正しく復号される。また、この方式において、

H(S|XY) = H(S|UY) = H(S|VX) = H(S|UV) = H(S)

であることが示される。H(・|・)は条件つきエントロピーを表す。すなわち、２つのシェア(X,Y)があっても鍵 K がないと S の情報は一切得られず、鍵の一部とシェアの一部があっても同様に S の情報は一切漏れない。

参加者１がシェアの改ざんに成功する確率を $P_{S,1}$, 参加者２がシェアの改ざんに成功する確率を $P_{S,2}$ と書く。我々は次の定理を得た。

定理４ ([13]) 上記のシェア生成・復号アルゴリズムでは

$$\max\{P_{S,1}, P_{S,2}\} \le \frac{q-2}{p-1} \tag{17}$$

が成り立つ。

式(17)より、q≪p を満たすように p を十分大きく選んでおけば、改ざんの成功確率は任意に小さくできる。従来研究にも改ざんの成功確率が式(17)の右辺の値になる方式が存在したが[6]、その方式は difference set と呼ばれる特別な組合せ構造を利用して Shamir のしきい値法を利用する方式になっており、いつも構成できるとは限らなかった、定理４で実質的な条件は p が素数（または素数のべき乗）であるという、有限体の存在条件だけであり、秘密分散法で符号器・復号器間で鍵が共有できるという条件のもとでは、緩やかな条件のもとで改ざんの成功確率の上界を得ることができる。定理４の証明は割愛するが、X'を参加者１が改ざんしたシェアとするとき、(S,Y). X, X' がこの順にマルコフ連鎖をなすことが鍵になる。

上記のシェア生成・復号アルゴリズムにおいては、なりすまし攻撃の成功確率も評価することができ、q/p 以下であることを容易に示すことができる。

上記の方式はシェアが m 個の場合に容易に拡張できる。例えばシェアが３個の場合は、K=(U,V,W)、３個のシェアを X, Y, Z として、シェア生成アルゴリズムを

(E1) E_1, E_2 を $E_1 + E_2 \neq S$ を満たすように S から一様ランダムに選ぶ。
(E2) $X = U^{-1} E_1$, $Y = V^{-1} E_2$, $Z = W^{-1} (S - E_1 - E_2)$ を出力する。

という形に拡張すればよい。一般に、シェアが m 個の場合であっても、定理４が成り立つことを示すことができる。

【参考文献】

[1] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, n¥ pp, 612—613, 1979.
[2] R. Blackley, "Safegarding cryptographic keys," *Proc. AFIPS 1979; National Computer Conference*, vol. 48, pp.313—317, 1979.
[3] R. J. McEliece and D. V. Sarwate, "On the secret sharing scheme and Reed-Solomon codes," *Communications of the ACM,* vol. 24, no. 9. pp. 583—584, 1981.
[4] M. Tompa and H. Woll, "How to share a secret with cheaters," *J. Cryptology*, no. 1, p. 133—138, 1988.
[5] M. Carpentieri, A. De Santis, and U. Vaccaro, "Size of shares and probability of cheating in threshold scheme," *Proc. Eurocrypt '93*, LNCS 765, Springer Verlag, pp. 118—125, 1994.
[6] W. Ogata, K. Kurosawa and D. R. Stinson, "Optimum secret sharing scheme secure against cheating," *SIAM J. Discrete Mathematics*, vol. 20, no. 1, pp. 79—85, 2006.
[7] E. D. Karnin, J. M. Greene, and M. E. Hellman, "On secret sharing systems," *IEEE Trans.. Inf. Theory*, vol. 29, pp. 35—41, 1983.
[8] H. Koga, M. Iwamoto and H. Yamamoto, "Coding theorems for a (2,2)-threshold scheme secure against impersonation by an opponent," *Proc. 2009 IEEE ITW*, pp.188—192,. Taomina, Italy, 2009.
[9] H. Koga, "Coding theorems for the threshold scheme for a general source," *IEEE Trans. Inf. Theory,* vol. 54, no. 6, pp. 2658—2677, 2008.
[10] T. S. Han, *Information-Spectrum Methods in Information Theory*, Springer-Verlag, 2003.
[11] H. Koga, Two generalizations of a coding theorem for a (2,2)-threshold scheme with a cheater,* to appear in *Proc. Int. Symp. Inf. Theory and its Appl*, Oct. 2010.
[12] S. Vembu, S. Verdu, and Y. Steinberg, "The source-channel separation theorem revisited," *IEEE Trans. Inf. Theory, vol. IT-41*, no. 1, pp. 44—54, 1995.
[13] 古賀, "共通の鍵をもつ秘密分散法の提案とそのシェアの改ざんに対する安全性," 2010 年暗号と情報セキュリティシンポジウム予稿集, IF2-3, 2010.

〈発 表 資 料〉

題　　名	掲載誌・学会名等	発表年月
Coding theorems for a (2,2)-threshold scheme secure against impersonation by an opponent	Proc. 2009 IEEE ITW, pp. 188- 192	2009 年 10 月
共通の鍵をもつ秘密分散法の提案とそのシェアの改ざんに対する安全性	2010 年暗号と情報セキュリティシンポジウム予稿集	2010 年 1 月
Two generalizations of a coding theorem for a (2,2)-threshold scheme with a cheater	Proc. 2010 ISITA (予定)	2010 年 10 月