情報処理学会論文誌 Vol.59 No.9 1570–1580 (Sep. 2018)

結託耐性符号を用いた名簿システムの分割攻撃への対応

木下 盾[1,a)] 上原 哲太郎[2,b)]

受付日 2017年11月27日, 採録日 2018年6月8日

概要：多くの企業が扱う顧客個人情報は事業活動のためには不可欠のものだが，同時に安全管理の責任が求められる．そのため，利活用できる状態でいかに漏洩を防止するかが課題である．特に内部不正による漏洩は，防止や検知が困難である．そこで，対策として従業員ごとにパターンを変えて名簿へダミー個人情報を埋め込む手法がある．しかし，これは結託攻撃によるダミー個人情報の除去が考えられるため，結託耐性符号の利用が必要である．また，名簿は分割される可能性が高く，分割によって結託耐性符号が除去されることも考えられる．本研究では，結託耐性符号を用いた名簿システムとその分割攻撃への対応手法を提案する．

キーワード：結託耐性符号，コンテンツ保護，個人情報保護，電子透かし，不正者追跡

A Countermeasure for Dividing-attack on the Address-list Managing System with ACC

Jun Kinoshita[1,a)] Tetsutaro Uehara[2,b)]

Received: November 27, 2017, Accepted: June 8, 2018

Abstract: Many of the business firms are utilizing the customer personal information for their business activities. At the same time, they are responsible to prevent leakage. It is a big issue for them to utilize the personal information keeping the safety to prevent the leakage. In particular, it is difficult to prevent ant detect the leakage due to internal fraud. Therefore, there is a method of embedding the dummy personal information to the roster by changing the pattern for each employee as a countermeasure. However, because of the threat of the removal of the dummy personal information by collusion attacks, it is necessary to use a collusion resistant code. Moreover, the customer list is likely to be divided into pieces, it is conceivable to collusion resistance code is removed by the division. In this paper, we propose a countermeasure as the list system using collusion code to the splitting attacks.

Keywords: anti-collusion code, contents protection, information leakage, digital watermark, traitor tracing

1. はじめに

情報化社会が進む今日，多くの企業は大量の顧客個人情報をデジタル化し，管理・分析を行っている．しかし，デジタル化された顧客名簿は複製や持ち出しが容易に行えるため，漏洩対策が必要となる．

特に，内部不正による名簿漏洩事件は事件1件あたりの漏洩する個人情報の件数が多い傾向 [1], [2] にある．さらに，内部不正は検知が難しい．一般的な内部不正の対策として，監視の強化やアクセス制限があげられるが，これらの対策は業務効率の低下を招きかねない．

業務効率の低下を招かない名簿の漏洩抑止としてダミー個人情報の埋め込みがある．ダミー個人情報とは，実在しない偽の個人情報である．これを，従業員がダウンロードする顧客名簿に埋め込む手法がある．名簿が名簿業者に売却され，他の企業に販売されると，ダミー個人情報の住所やメールアドレスにダイレクトメール（以下 DM）や電子メールが届き，自社からの名簿漏洩を検知することができ

[1] 立命館大学大学院情報理工学研究科
Graduate school of information science and engineering, Ritsumeikan University, Kusatsu, Shiga 525–8577, Japan
[2] 立命館大学情報理工学部
College of information science and engineering, Ritsumeikan University, Kusatsu, Shiga 525–8577, Japan
a) kinoshita@cysec.cs.ritsumei.ac.jp
b) t-uehara@fc.ritsumei.ac.jp

る．さらに，名簿をダウンロードした従業員や日時によって，異なるダミーの個人情報を埋め込むことで，漏洩日時や関与した従業員を特定することができる．

しかし，この手法は結託攻撃を行われると，ダミー個人情報が除去され攻撃者の特定が不可能になる．結託攻撃は電子透かしへの攻撃手法であり，複数の情報を比較し，異なる部分を除去する攻撃である．これの対策として，結託攻撃に耐性を持つ結託耐性符号の利用が考えられる [3]．

結託耐性符号を利用しても，名簿ならではの問題がある．名簿の特性上，分割や抜き出し等を行われる可能性が高い [4]．その結果，ダミー個人情報も除去されてしまう．これらを分割攻撃と呼ぶ．

ダミー個人情報の埋め込みは，除去することが困難なダミー個人情報が埋め込まれていることを組織内部の人間に周知することで，顧客名簿漏洩の心理的抑止力になる．本研究では，結託耐性符号を用いた名簿システムの提案とその分割攻撃への対応を提案する．

2. 研究背景

2.1 内部不正の分類

Cappli らは，内部不正を下記の 3 つに分類した [5]．

IT Sabotage

従業員が，個人や企業への復讐を目的に，社内の重要な情報やシステムの破壊を行う．

Theft of intellectual property

退職する従業員が，転職を有利に進めることを目的に，業務上知り得た秘密や知的財産を漏洩する．

Fraud

従業員が，金銭目的で，個人情報等の社内の情報を入手，漏洩する．

内部不正として狙われる情報として，知的財産と顧客名簿があげられる．知的財産は企業の特許や研究等が含まれ，産業スパイや転職によって競合他社に狙われる傾向にある．一方で，顧客名簿は金銭目的で狙われる傾向にある．これは，顧客名簿は知的財産等とは異なり，名簿業者 [4] の存在により売却しやすく，漏洩した名簿から漏洩元の企業を特定するのも困難なためである．

2.2 内部不正による顧客名簿漏洩

本研究では内部不正による顧客名簿漏洩に着目する．攻撃者は名簿をダウンロードする権限を有する従業員である．本研究で，顧客名簿と表記した場合は企業が持つ全顧客の名簿とする．単に名簿と表記した場合は，権限を有する従業員が顧客名簿の一部をダウンロードしたものとする．攻撃者は金銭目的のため，大量の顧客個人情報を含む名簿を名簿業者に売却するものとする．

2.3 内部不正対策

独立行政法人情報処理推進機構が行った調査 [6] によれば，内部不正経験者は，アクセスログの監視やアクセス権限の制限，監視体制の強化，罰則規定の強化等によって内部不正への動機が低下するという結果が出ている．しかし，経営者，システム管理者は，監視体制の強化や罰則規定の強化が内部不正への動機の低下につながると考えている割合は少ないという結果が出ている．内部不正経験者と経営者の認識の相違は問題であり，監視体制の強化や罰則規定の強化につながる対策を講じる必要がある．

独立行政法人情報処理推進機構が発表した組織における内部不正防止ガイドライン [7] では，以下の 5 つを内部不正防止の基本 5 原則と定義している．

- 犯行を難しくする．
- 捕まるリスクを高める．
- 犯行の見返りを減らす．
- 犯行の要因を減らす．
- 犯罪の弁明をさせない．

内部不正にはこれらの原則に従った対策が効果的である．本論文では，このうち「捕まるリスクを高める」ことに着目した対策を提案している．

2.4 ダミー個人情報埋め込み

ダミー個人情報の埋め込みは，実在しない偽の個人情報（ダミー個人情報）を名簿に埋め込む手法である．内部不正防止ガイドラインの基本 5 原則では，「捕まるリスクを高める」にあてはまる．これは，ダミー個人情報が埋め込まれていることを組織内部の人間に周知することで，顧客名簿漏洩の心理的抑止力になる．また，ダミー個人情報を含んだ顧客名簿が漏洩しそれが DM や電子メールの送付先として使用された場合，企業は名簿の漏洩を検知することができる．さらに，漏洩した名簿がどの名簿業者を経由し，他企業へ販売されたか追跡できる．

ここで，どれがダミー個人情報であるか利用者（攻撃者）が気づかなければ，ダミー個人情報を取り除かれることはない．ダミー個人情報の名前や年齢，性別はランダムで生成される．住所や電話番号，メールアドレスは監視されたものを利用する．すなわち，ダミー個人情報は実在する人物と一致する可能性はなく，攻撃者は個人情報を見ただけでダミー個人情報と判別することはできない．本研究でも，個人情報を見ただけではダミー個人情報と判定できないと仮定する．

このような考え方は電子透かしのコンテンツ追跡やソーシャル DRM と類似している．ある情報（画像，動画，文書，名簿等のコンテンツ）をカバーデータとし，そこに識別情報（著作権，購入者情報）を埋め込む．電子透かしにおいて，識別情報を埋め込む際には以下の 3 点が重要となる [8]．

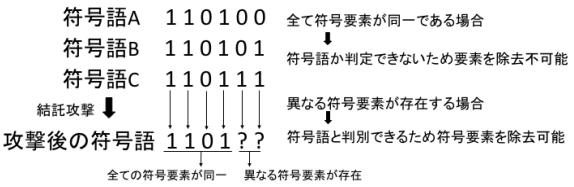

図 1 Marking Assumption を満たした結託攻撃
Fig. 1 Collusion attack under Marking Assumption.

- 識別情報を埋め込まれても，カバーデータの品質を損なわない．
- カバーデータの編集や圧縮等，各種改変が行われても識別情報が消えない．
- 識別情報の読み書きは許可されたもののみが可能であること．

ダミー個人情報の埋め込みが一般的な電子透かしと異なる点は，識別情報の埋め込みによりカバーデータの内容が変化する点である．名簿にダミー個人情報が加わることで別の名簿になってしまう．しかし，個人情報は年月とともに引っ越しや出産，結婚等様々な要因で変更される．その変更が必ず顧客名簿に反映されるとは考え難い．このような点から，ダミー個人情報が数パーセント含まれていても名簿としての利用価値は大きく損なわれることはないと考えられる．また，顧客の同意のもと，DM や電子メールの送信を目的に顧客名簿を協力会社等に渡す場合，ダミー個人情報の存在は問題にならない．

攻撃者を検挙するまでのプロセスが，電子透かしよりも確立されていることもダミー個人情報埋め込みの特徴である．電子透かしは，埋め込まれたコンテンツが悪用（複製，配布）されても，それ自体がコンテンツの権利を持つ企業にわたらない限り，攻撃者を特定することはできない．一方で，ダミー個人情報埋め込みは，個人情報が悪用（DM や電子メールを送付）されることで，攻撃者を特定し，検挙につながる．

2.5 結託耐性符号

電子透かしは，攻撃者により識別情報を改ざん，破壊される可能性がある．電子透かしへの攻撃には，カバーデータの数によって単独攻撃と結託攻撃に分類することができる．単独攻撃とは，カバーデータを編集，切り取り，圧縮等によって識別情報を破壊する攻撃である．結託攻撃とは，複数のカバーデータを比較し，異なる部分を改ざん，削除することで，識別情報を除去する攻撃である．ここで，比較するカバーデータの数を結託数と呼ぶ．

結託攻撃に耐性を持つ符号として結託耐性符号がある．この符号は，Marking Assumption [9] の成立（**図 1**）を前提に，結託数が一定数までであれば，結託攻撃に加担した識別情報を少なくとも 1 人は特定することができる．結託耐性符号を利用することで，結託攻撃を行われても除去されない識別情報を埋め込むことができる．Marking Assumption とは，攻撃者が結託攻撃を行う際に，符号の要素が異なる部分しか改ざんできないという仮定である．攻撃者が検出した要素に対し操作できる範囲によって仮定の強さが異なる（**表 1**）．ただし，2 元符号の場合は，expanded narrow-sense と wide-sense は一致する．

表 1 Marking Assumption
Table 1 Marking Assumption.

仮定	検出した要素への操作	仮定の強さ
narrow-sense	攻撃者の符号要素	弱い
expanded narrow-sense	攻撃者の符号要素か除去	強い
wide-sense	任意の要素	強い
expanded wide-sense	任意の要素か除去	最も強い

結託攻撃によって一部が破壊された符号から，攻撃者の符号を特定することを追跡と呼ぶ．追跡が可能な結託耐性符号として，SEC（c-secure code with ε-error）符号 [10]，IPP（c-identifiable parent property code）符号 [11]，TA（c-traceability）符号 [12]，Tardos 符号 [13]，ACC（Anti-Collusion Code）符号 [14] があげられる．SEC 符号は，弱い Marking Assumption で定義されおり，追跡能力が低い．また，IPP 符号は他の符号よりも結託耐性が低い [15]．

Tardos 符号は小さい符号長で符号語を生成でき，攻撃者を特定できる可能性は高く攻撃者でない者を攻撃者と間違える確率が低いが，保証はない．TA 符号や ACC 符号は結託数が一定以下であれば，必ず攻撃者を特定可能であり，攻撃者でない者を攻撃者と間違う可能性がないことを保証している．

一般的に多元符号は，符号要素を交換するシンボル選択攻撃という結託攻撃を想定し，2 元符号は複数の符号の平均値をとる平均化攻撃という結託攻撃を想定している．TA 符号は多元符号のため複数の符号要素により符号を構成されており，符号語を符号要素以外に書き換えられた場合に攻撃者を特定できなくなる．一方，ACC 符号は平均化による結託攻撃を想定しており，符号要素以外への符号語の改竄を許容している．

本研究では，符号要素が消失シンボルとなる可能性があるため，追跡性を持ち，符号要素以外への符号の改竄を許容している ACC 符号を利用する．

2.6 ACC 符号

ACC 符号は BIBD（Balanced Incomplete Block Design）と呼ばれる集合によって構成され，利用する BIBD によって符号の性質が異なる．

2.6.1 BIBD

BIBD は，複数の集合を作り出し，それらの集合の間にバランスを持たせたものである Block Design の一種である．具体的には，点と呼ばれる有限集合 X と，ブロックと

表 2 BIBD のパラメータと ACC 符号のパラメータ
Table 2 Parameters for BIBD and ACC code.

BIBD のパラメータ	ACC 符号
v：点 x の元の個数	符号長
b：ブロック B の個数	符号数
r：ある点を含むブロックの個数	0 が持つ情報
k：1 つのブロックに含まれる点の個数	符号の 0 の数
λ：異なる 2 つの点を含むブロックの個数	2 つの 0 が持つ情報

$BIBD(7,3,1)$
$B = \{ B_0(0,1,3), B_1(1,2,4), B_2(2,3,5), B_3(3,4,6), B_4(0,4,5), B_5(1,5,6), B_6(0,2,6) \}$

	B0	B1	B2	B3	B4	B5	B6
0	0	1	1	1	0	1	0
1	0	0	1	1	1	0	1
2	1	0	0	1	1	1	0
3	0	1	0	0	1	1	1
4	1	0	1	0	0	1	1
5	1	1	0	1	0	0	1
6	1	1	1	0	1	0	0

図 2 BIBD の ACC 符号化
Fig. 2 Making an ACC code from a BIBD.

呼ばれる X の部分集合族 B の組 (X, B) である．部分集合族 B は b 個のブロックで構成され，1 つのブロックには k 個の点 x が含まれる．

$$X = \{x_0, x_1, x_2, \ldots, x_{v-1}\}$$
$$B = \{B_0, B_1, B_2, \ldots, B_{b-1}\}$$

さらに以下の 2 つの性質（均衡性）を持つ．

- 任意の点 x に対し，x を含むブロックは r 個存在する．
- 任意の異なる 2 つの点をともに含むブロックは λ 個存在する．

各パラメータは以下の関係を持ち [8]，BIBD は $BIBD(v, b, r, k, \lambda)$ または，$BIBD(v, k, \lambda)$ と表される．

$$vr = bk$$
$$\lambda v\,(v-1) = bk\,(k-1)$$

2.6.2 符号化

ACC 符号はバイナリで構成され，符号要素の大部分を 1 が占める．BIBD のブロック内の要素は，ACC 符号の要素が 0 である部分を示している．ACC 符号の符号長は BIBD の有限集合 X の点の個数 v となり，X の部分集合族 B のブロックの個数 b が生成できる符号の数となる．1 つのブロックの要素数 k は ACC 符号 1 つに存在する 0 の数となる．λ は異なる 2 つの 0 を含む ACC 符号の数を示している．以上の BIBD のパラメータと ACC 符号のパラメータの関係を**表 2** にまとめる．

例として BIBD$(7,3,1)$ の ACC 符号化を**図 2** に示す．BIBD の各ブロックは ACC 符号の各符号語を意味し，点はその符号語の符号要素が 0 であることを示している．図 2 の場合，$B_0(0,1,3)$ は符号語の 0，1，3 ビット目の符号要素が 0 であることを示すため，生成される符号語は 0010111 となる．

表 3 代表的な BIBD
Table 3 Typical BIBDs.

BIBD	(v, k, λ)
Fano plane	$(7, 3, 1)$
Hadamard design	$(4m+3, 2m+1, m)$
Projective plane	$(m^2+m+1, m+1, 1)$
Steiner triple system	$(v, 3, 1)$
affine plane	$(m^2, m, 1)$
unital	$(p^3+1, p+1, 1)$

2.6.3 検討

表 3 に代表的な BIBD を示す．表 3 から，アフィン平面（affine plane）を利用した BIBD と単位環（unital）を利用した BIBD が他よりも小さな符号長で，結託耐数や生成できる符号の数（以下符号数）が大きく，性能が良いと評価されている．これらは，Hou らが構築した [16] ACC 符号である．どちらも最も強い Marking Assumption（expanded wide-sense）を満たしている．アフィン平面による ACC 符号は，結託人数が多い結託攻撃にも耐性を持つが，符号数が少ない．単位環による ACC 符号は，結託数ではアフィン平面による ACC 符号に劣るが，符号をより多く生成できる．

結託耐性符号をダミー個人情報埋め込みに利用すると，符号長が短いほど，ダミー個人情報の埋め込み量が減り，符号数が多いほど，従業員およびダウンロード日時の識別数が増える．よって本研究では，同じ符号長でより多くの符号を生成できる，単位環による ACC 符号を利用する．

2.6.4 単位環による ACC 符号

任意の素数 p による単位環を利用した BIBD で構成される ACC 符号は符号長 p^3+1 bit で，$p^2(p^2-p+1)$ 個までの符号を生成でき，p 人までの結託を特定（以下，結託耐数）できる．

たとえば，$p=13$ の ACC 符号は，符号長が 2,198 bit で符号数が 26,533 個，結託耐数は 13 である．通常，26,533 個の符号数は符号長を 15 bit で表すことができるため，ACC 符号の符号語は大幅に冗長化されていることが分かる．そのため，符号語は符号要素がある程度欠落しても，識別可能である．

また，ACC 符号は $\lambda=1$ であるため，0 の符号要素が 2 つ決まれば，符号語は一意に決まる．また，0 の符号要素が 1 つ決まれば，全符号語 26,533 個の中から 182 個に絞り込むことができる．対して，1 の符号要素が 1 つ決まっても，26,533 個の符号の中から 26,351 個にしか絞り込むことができない．本研究では，1 つ符号要素で排除できる符号語の数のことを，符号要素が持つ情報量と定義する．符号要素の持つ情報量が大きいほど，符号の中から候補の符号語を絞り込むことができる．すなわち，0 の符号要素は 1 の符号要素に比べて大きな情報量を持っている．

2.7 関連研究

関連研究として，結託耐性符号の構成方法や性能を向上させる研究 [14], [16]，ランダム誤りへの耐性を高める研究 [17], [18] があげられる．また，結託耐性符号を利用したコンテンツ保護 [19] の研究も行われている．

名簿の漏洩防止の観点では，ダミー個人情報を埋め込む手法の特許 [20], [21], [22] やサービス [23] が存在する．ただし，これらは結託攻撃を考慮していない．

3. 研究手法

3.1 シナリオ

本論文で想定する名簿の漏洩の流れを図 **3** に示す．

(1) 複数の攻撃者がダミー個人情報が埋め込まれた名簿をダウンロードする．
(2) 複数の名簿を比較して異なる部分を除去する結託攻撃を行う．
(3) 結託攻撃後の名簿を名簿業者に売却する．
(4) 名簿業者は購入した名簿を他の名簿と結合させる．
(5) 名簿業者は名簿の一部を他の企業に販売する．
(6) 名簿を購入した企業はその名簿をもとに DM を送付する．
(7) ダミー個人情報の宛先に DM が届く．

3.2 ダミー個人情報埋め込み手法

ACC 符号の符号語をダミー個人情報の埋め込みパターンで表現する．符号語の各符号要素に 0 を意味するダミー個人情報と 1 を意味するダミー個人情報を割り当てる．そして，各符号要素（0 または 1）によって，埋め込むダミー個人情報を決定する．ここで，符号語の符号長を l とするとダミー個人情報は $l \times 2$ 件必要となる．

ダミー個人情報の埋め込み例を図 **4** に示す．図 4 では，符号語 $\alpha(1101)$ と $\beta(0111)$ が用意されている．各符号要素 $(0,1)$ には 1 ビット目から順番に，(鈴木，宮元)，(斎藤，田中)，(山崎，足立)，(小西，山本) というダミー個人情報が割り当てられている．(鈴木，斎藤，山崎，小西) は 0，(宮元，田中，足立，山本) は 1 を意味する．符号語 $\alpha(1101)$ は，

1 ビット目が 1 であるため，宮元
2 ビット目が 1 であるため，田中
3 ビット目が 0 であるため，山崎
4 ビット目が 1 であるため，山本

となり，(宮元，田中，山崎，山本) の 4 つのダミー個人情報で表すことができる．

名簿がダウンロードされた際，符号語に，ダウンロードしたユーザの情報を紐付け，データベースへ格納する．これは，ログの記録にもなる．符号語にユーザと日時を紐付けることで，攻撃者だけでなく，日時まで特定することが可能になる．一方で，同じユーザが異なる日時に名簿をダウンロードしても，異なる符号語が埋め込まれるため，攻撃者は 1 人で結託攻撃が可能になる．

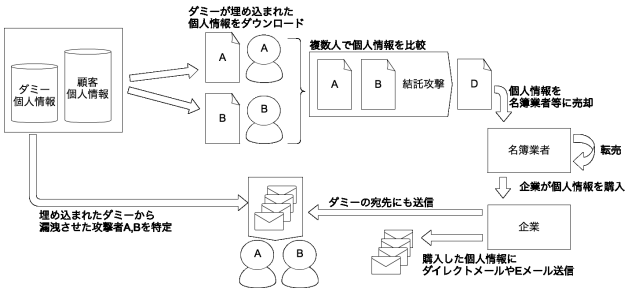

図 **3** 漏洩した名簿の流れ
Fig. 3 Flow of the leaked address list.

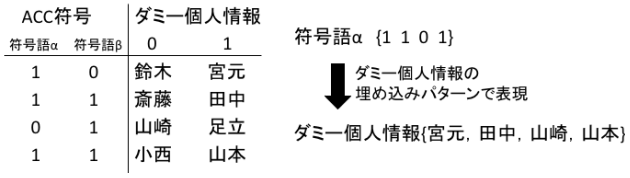

図 **4** 埋め込み手法
Fig. 4 The method to embed codes to an address list.

3.3 名簿へのシャッフル攻撃

名簿は名前や住所，年齢等によって並べ替えられる可能性が高い．この際，ダミー個人情報がシャッフルされた状態となる．しかし，ダミー個人情報には符号要素の位置（何ビット目か）および内容 $(0,1)$ が割り当てられているため，シャッフルに対して耐性を持つ．

3.4 名簿への結託攻撃

名簿への結託攻撃は複数の名簿を比較し，すべての名簿に存在しない個人情報を除去する攻撃である．このような結託攻撃に対して，ACC 符号は耐性を持つ．

ACC 符号は 2 元符号であるが，符号要素をダミー個人情報に対応させることで，名簿への埋め込みが可能になったと同時に，多次元符号になっている．よって，シンボル選択による結託攻撃の耐性を考える必要がある．シンボル選択による攻撃は，符号要素を他の符号要素と交換することで，識別情報を破壊，偽造する攻撃である．しかし，攻撃者は名簿からダミー個人情報のみを識別することができない．よって，ダミー個人情報はシンボル選択による攻撃にも耐性を持つ．

3.5 名簿への分割攻撃

ダミー個人情報が埋め込まれた名簿の分割を図 **5** に示す．ダミー個人情報は名簿全体に複数件埋め込まれ，埋め込まれたパターンで識別情報を表している．しかし名簿業者は名簿を複製，結合および分割を行うと考えられる．ダミー個人情報が含まれた名簿を複数の名簿に分割すること

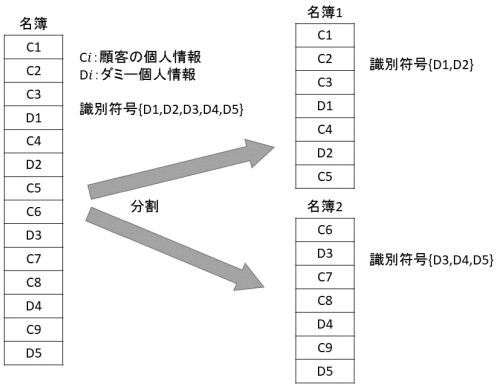

図 5 名簿の分割
Fig. 5 Division of an address list.

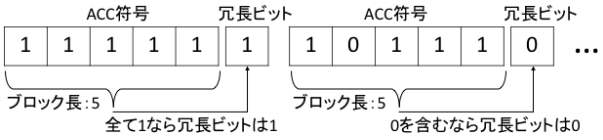

図 6 冗長ビット
Fig. 6 Redundant bits.

は，Marking Assumption を満たさない強い攻撃である．これを分割攻撃と呼び，分割攻撃への耐性を分割耐性と呼ぶ．また，規則性がなくダミー個人情報が除去されるため，符号の要素がランダムに欠落する攻撃ともとらえることができる．

名簿を分割をする際，属性（性別，住所，年齢等）によって分割する可能性がある．このため，ダミー個人情報の属性は，顧客名簿全体の分布を考慮して作成する必要がある．

3.6 冗長 ACC 符号

ACC 符号の符号語に，冗長にビットを付加することで，効率良く分割耐性を向上させることができる．本研究ではこれを冗長ビットと定義し，冗長ビットを付加した ACC 符号を冗長 ACC 符号と呼ぶ．冗長ビットは，符号を一定の長さでブロックとし，そのブロックの論理積をとったものである．ブロック長 5 の冗長ビットの例を**図 6** に示す．

ブロック長によって冗長ビットの性質は変化する．

ブロック長 = 1
冗長ビットは元の符号語と同値になり，符号語を冗長に 2 回埋め込むことを意味する．

ブロック長を小さくする
冗長ビットは 1 が多くなる．冗長ビットは ACC 符号の符号要素と同様に，結託耐性が強くなる．

ブロック長を大きくする
冗長ビットは 0 が多くなる．冗長ビットが持つ情報量が大きくなり，冗長ビット 1 ビットで復元できるビットが増える．

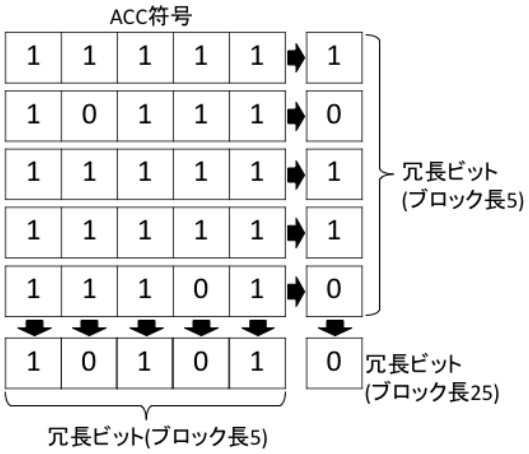

図 7 2 次元冗長ビット
Fig. 7 2-dimensional redundant bits.

ブロック長 = 符号長
冗長ビットはすべて 0 となる．

符号要素が消失しても，冗長ビット 1 であれば，対応するブロックが含む符号要素すべてを（すべて 1）を復元することができる．冗長ビットが 0 の場合は，対応するブロックが含む符号要素を復元することはできないが，候補の符号語をいくつかに絞り込むことができる．これは 0 が持つ情報量が大きいためである．

冗長ビットによる誤り訂正は，従来の誤り訂正と性質が異なる．通常，誤り訂正は符号語の一部に誤りが発生することを想定し，符号を完璧に復元することを目的としている．冗長ビットは，符号語の大部分が消失することを想定している．そして，符号要素の大部分が欠落した状態から少しでも復元することを目的としている．これは，ACC 符号が持つ誤り訂正能力と組み合わせることで効果が期待できる．

また，**図 7** に示すように 2 次元に冗長ビットを付加することもできる．この 2 次元冗長ビット付加において，右下のビットは複数の冗長ビットの論理積をとったビットであり，ブロック長は他の冗長ビットのブロック長の 2 倍である．すなわち，2 次元冗長ビットによる拡張は，同じブロック長の冗長ビットを 2 回付加し，さらに 2 倍のブロック長の冗長ビットを付加したことになる．

4. シミュレーション実験

計算機を用いて，冗長 ACC 符号の性能を実験により測定する．プログラムには C 言語を使用し，疑似乱数の生成にはメルセンヌ・ツイスタを用いた．

4.1 実験に用いる符号

本実験では，$p = 13$，$BIBD(v, b, k, r, \lambda) = (2198, 26533, 14, 169, 1)$ の単位環を用いた BIBD による ACC 符号を使用する．BIBD の生成には数式処理システム MAGMA [25]

を用いた．

表 4 に ACC 符号のパラメータ，**表 5** に各ブロック長の冗長 ACC 符号のパラメータを示す．冗長 ACC 符号の各符号語の符号長は 3,000 bit であり，識別数と結託耐数は拡張前の ACC 符号の符号語と同等である．冗長ビットを付加しても符号長が 3,000 bit に満たない場合は，符号長が 3,000 bit になるまで同じ冗長ビットを冗長に付加している．ブロック長が 1 の場合は，符号長が 4,396 bit 必要になる．しかしこれは ACC 符号を冗長に付加しているだけであるため，差分の 1,396 bit は削除した．

ACC 符号は結託耐数を超える結託攻撃や，Marking Assumption を満たさない攻撃によって大幅に符号要素が欠落すると，攻撃者を特定できなくなる．この場合，攻撃者以外の符号語も攻撃者の候補の符号語として検知する．この検知された符号語を攻撃者候補符号語，攻撃者候補符号語の数を攻撃者候補数，攻撃者の符号語を攻撃者符号語と定義する．攻撃者候補符号語の中には攻撃者符号語が必ず含まれる．ACC 符号を A，攻撃者候補符号を B，攻撃者符号を C とすると A，B，C は式 (1) の関係を持つ．

$$A \supseteq B \supseteq C \tag{1}$$

また，攻撃者候補符号語のうち，攻撃者以外の符号語を不当候補符号語，不当候補符号語の数を不当候補数と定義する．攻撃者候補数を X，攻撃者数を Y，不当候補数を Z とすると X，Y，Z は式 (2) の関係を持つ．

$$X = Y + Z \tag{2}$$

4.2 結託攻撃実験

冗長 ACC 符号の結託耐性を検証するため，ACC 符号と各ブロック長の冗長 ACC 符号に対してランダムな結託数で結託攻撃を行い，攻撃後の符号後に対して不当候補数を測定する．この実験をランダムな結託数と各ブロック長で 10,000 回ずつ行う．本実験では，ランダムに結託数個の ACC 符号語を選択し，それらを比較し，1 つでも異なる符号要素がある場合は，その符号要素を消失シンボルとすることで，結託攻撃としている．

表 4 ACC 符号 $p = 13$ のパラメータ
Table 4 The parameters for the ACC code with $p = 13$.

符号長	識別数	結託耐数	要素が 0 の数
2,198 bit	26,533	13	14

表 5 冗長 ACC 符号のパラメータ
Table 5 The parameters of the redundant ACC code.

ブロック長	平均ハミング重み	符号長
1	2,980.98	
10	2,935.57	3,000 bit
100	2,591.03	
1000	2,252.18	

また，結託攻撃後の冗長ビットが除去された割合を求めるため，各ブロック長の冗長 ACC 符号に対して結託攻撃（結託数：2, 5, 8, ..., 47）を行う．攻撃後の冗長 ACC 符号のうち，削除された冗長ビットを測定し，付加した冗長ビットが削除された割合を求める．この実験を各結託数と各ブロック長で 200 回ずつ行う．

4.3 分割攻撃実験

冗長 ACC 符号の分割耐性を検証するため，ACC 符号と冗長 ACC 符号の符号要素をランダムに欠落させ，不当候補数を測定する．これを ACC 符号と各ブロック長の冗長 ACC 符号で行う．本実験では，符号要素を一定数ごとランダムに欠落させることで，Marking Assumption を満たさない攻撃となり，これを分割攻撃と見なす．欠落ビット数は 0 か符号長である 3,000 までランダムに決定し，各ブロック長で 10,000 回ずつ実験を行う．

4.4 結託分割攻撃実験

冗長 ACC 符号の結託耐性と分割耐性同時に検証するため，ACC 符号と各ブロック長の冗長 ACC 符号に対してランダムな結託数で結託攻撃を行う．その後，攻撃後の ACC 符号と冗長 ACC 符号の符号要素を一定数ごとランダムに欠落させ，不当候補数を測定する．これを結託分割攻撃と呼ぶ．これを ACC 符号と各ブロック長の冗長 ACC 符号，各結託数で行う．欠落ビット数は 0 か符号長である 3,000 までランダムに決定し，各ブロック長で 10,000 回ずつ実験を行う．

5. 実験結果

結託耐性の実験は ACC 符号と各ブロック長の冗長 ACC 符号に対して行ったが，すべて同じ結果となった．その結果を**図 8** に示す．縦軸は，不当候補数を示し，横軸は結託数を示す．結託攻撃後の冗長ビットが除去された割合を**図 9** に示す．縦軸は結託攻撃によって除去されてない冗長ビットの割合を示し，横軸は結託数を示している．

分割攻撃実験結果を各ブロック長ごとにグラフに示す．

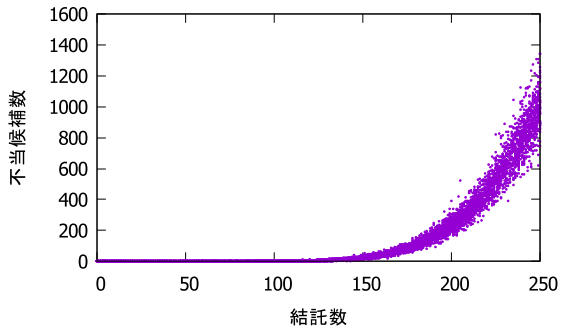

図 8 結託耐性実験結果
Fig. 8 Result of an experimental test for collision resistance.

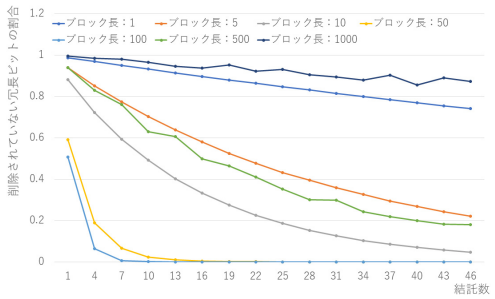

図 9 実験結果：冗長ビットの結託耐性
Fig. 9 Collision resistance of redundant bits.

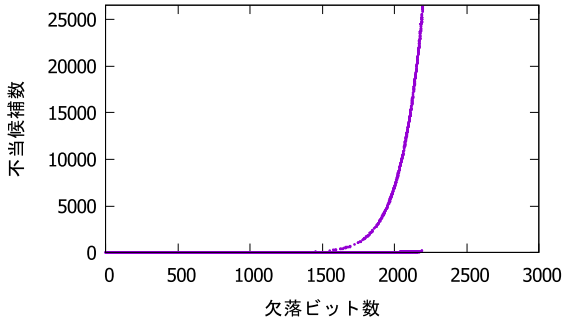

図 10 分割攻撃実験（ACC 符号（$p = 13$））
Fig. 10 Division attack test (ACC code ($p = 13$)).

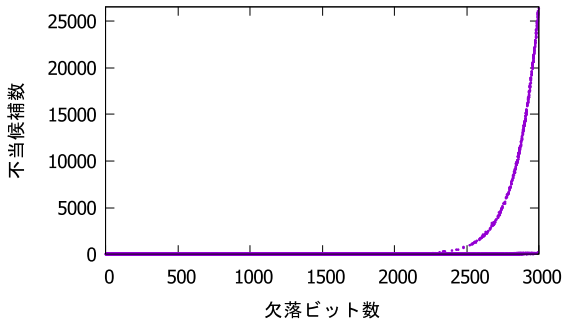

図 11 分割攻撃実験（ブロック長：1）
Fig. 11 Division attack test (block size: 1).

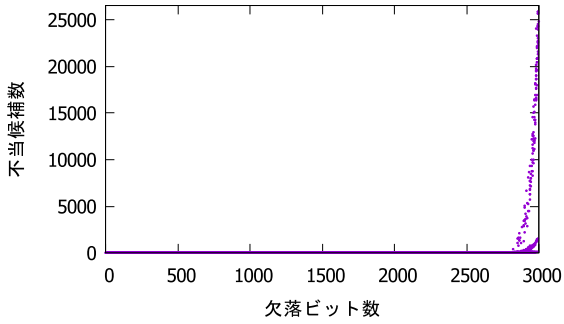

図 12 分割攻撃実験（ブロック長：10）
Fig. 12 Division attack test (block size: 10).

ACC 符号の結果を図 **10** に，ブロック長 1 の結果を図 **11** に，ブロック長 10 の結果を図 **12** に，ブロック長 100 の結果を図 **13** に，ブロック長 1,000 の結果を図 **14** に示す．縦軸は，不当候補数を表し，横軸は欠落ビット数を表す．

結託分割攻撃実験結果を各ブロック長ごとにグラフに示す．ACC 符号の結果を図 **15** に，ブロック長 1 の結果を図 **16** に，ブロック長 10 の結果を図 **17** に，ブロック長 100 の結果を図 **18** に，ブロック長 1,000 の結果を図 **19** に

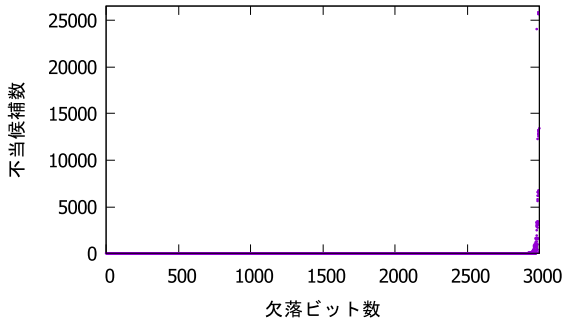

図 13 分割攻撃実験（ブロック長：100）
Fig. 13 Division attack test (block size: 100).

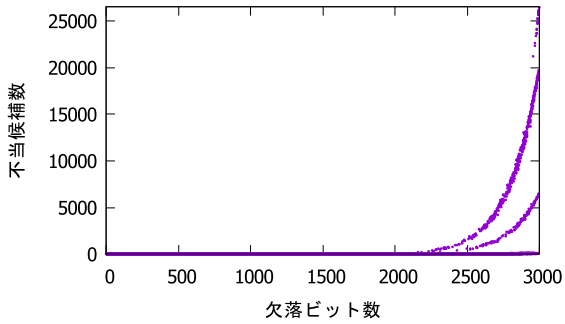

図 14 分割攻撃実験（ブロック長：1,000）
Fig. 14 Division attack test (block size: 1,000).

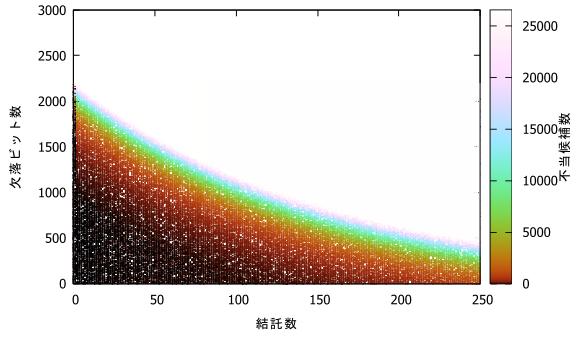

図 15 結託分割攻撃実験（ACC 符号（$p = 13$））
Fig. 15 Collision and division attack test (ACC code ($p = 13$)).

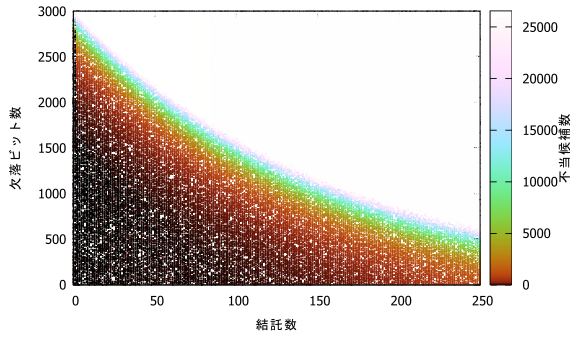

図 16 結託分割攻撃実験（ブロック長：1）
Fig. 16 Collision and division attack test (block size: 1).

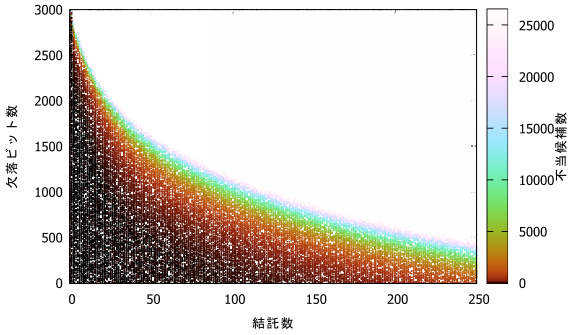

図 17 結託分割攻撃実験（ブロック長：10）
Fig. 17 Collision and division attack test (block size: 10).

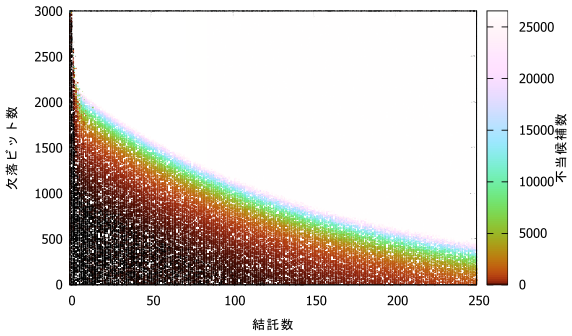

図 18 結託分割攻撃実験（ブロック長：100）
Fig. 18 Collision and division attack test (block size: 100).

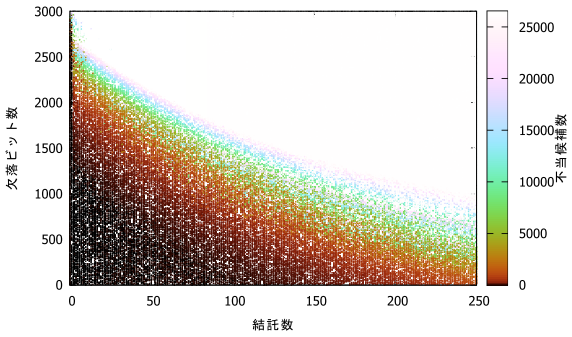

図 19 結託分割攻撃実験（ブロック長：1,000）
Fig. 19 Collision and division attack test (block size: 1,000).

示す．縦軸は，欠落ビット数を表し，横軸は結託数を表す．各点の色の濃淡は，不当候補数を表す．

6. 評価・考察

本実験では，結託攻撃後も不当候補数がより小さくなる符号を選択することが望ましく，これが小さいほど性能が良いとする．結託攻撃実験結果はすべてのグラフが類似していた．また，結託攻撃実験を各ブロック長の符号に対して，同じ種の擬似乱数で行うと，まったく同じ結果図 8 が得られた．これは，冗長ビットの結託耐性は ACC 符号の結託耐性よりも強くならないためである．すなわち，冗長ビットのブロック長が結託耐性に影響を与えないこと意味している．ただし，図 9 より，冗長ビットが削除される割合は，ブロック長 100 が高いことが分かる．

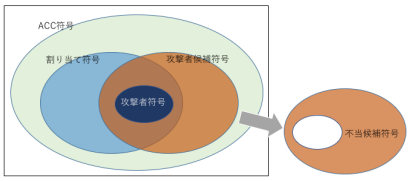

図 20 候補符号と割当て符号
Fig. 20 Candidate codes and assigned codes.

分割攻撃実験結果図 11〜14 より，すべてのグラフにおいて，欠落ビットが 3,000 近くでも不当候補数が 0 に近い場合がある．これは，符号要素の大部分が消失しても，符号を一意となることを意味している．特に，ブロック長 100（図 13）が欠落ビットが大きい場合の性能が良いことが分かる．これらより，分割耐性において，ブロック長は大きすぎても小さすぎても良くないことが分かる．ブロック長 100 以外のグラフでは，欠落ビット数 2,500 前後から不当候補数が大きくなる場合がある．特にブロック長 100 のグラフ（図 14）では，不当候補数が 3 パターンに分類できる．これは，ACC 符号要素は 2 つの 0 の位置が分かれば，符号語が一意に決まる性質と関係している．3 パターンのうち，不当候補数が 0 に近い値は符号語に 0 が 2 つ以上含まれていた場合であり，不当候補数が最も多い値，符号語に 0 が含まれていない場合であり，どちらでもない中間の値は 0 を 1 が 1 つ含む場合であると考えられる．

結託分割攻撃実験結果図 16〜19 より，ブロック長が小さい方が性能が安定していることが分かる．ブロック長が大きくなると図 19 のように色が分散する．色の分散は，結託数と欠落ビット数が似ている状況であっても，結果が大きく異なることを意味している．これは，結託攻撃に加え，ブロック長 1,000 はハミング重みが小さい（表 5）ため，符号語に 0 が含まれていない場合が増加し，性能が安定しないと考えられる．ブロック長 10，100（図 17，図 18）は結託数が小さい場合に限り，ブロック長 1（図 16）よりも，不当候補数が少なく，性能が良いといえる．しかし，結託数が大きい場合は不当候補数が多く，性能が悪い．これは，冗長ビットの多くが結託攻撃で削除されたためだと考えられる．

本研究実験では，全符号語 26,533 個から攻撃者候補数を測定した．しかし，実際には全符号語のうち，名簿をダウンロードした情報が紐付いている符号語の中から攻撃者候補数を測定することになる．ここで，名簿をダウンロードした情報が紐付いている符号語の集合を割当て符号と定義すると，ACC 符号と攻撃者候補符号と割当て符号と攻撃者符号の関係は**図 20** で表すことができる．結託攻撃後も攻撃者候補数がより小さくなる符号を選択することが望ましい．

そして，攻撃者符号は，必ず攻撃者候補符号と割当て符号の共通部分に存在する．割当て符号数によって，攻撃者

候補数が大幅に小さくなることも考えられる．

冗長 ACC 符号は，ACC 符号のパラメータとブロック長を変えることで，性能が変化する．名簿の重要性や使用目的，ダウンロードされる頻度や，結託攻撃が行われるリスクによって，各パラメータを調整し，状況に適切な性能の冗長 ACC 符号を生成することが可能である．

ダミー個人情報埋め込みでは，ダミー個人情報の住所や電話番号，メールアドレス等を維持・監視するコストがかかる．これらは，顧客名簿を持つ企業にとって大きな負担になる．しかし，ダミー個人情報の名前や年齢，性別と連絡先の組合せを変えることで，住所や電話番号，メールアドレスは使いまわすことが可能である．そのため，ダミー個人情報埋め込みを漏洩監視サービスとして提供することも考えられる．

ダミー個人情報が名簿に存在しているだけでは，法的に攻撃者を断定することは困難であると考えられる．しかし，名簿システムのログや名簿業者の記録等とともに，証拠の 1 つとなる．

7. まとめと今後の課題

本研究では，結託耐性符号を用いたダミー個人情報の埋め込み手法と分割耐性向上のための結託耐性符号の冗長ビット付加手法を提案した．冗長ビット付加した ACC 符号を用いてダミー個人情報を名簿に埋め込むことで，名簿は結託耐性と分割耐性を持ち，追跡可能性が向上する．また，冗長ビットと ACC 符号のパラメータを調整することで，冗長 ACC 符号の性質を想定される状況に合わせることが可能となった．これらを，従業員に周知させることで，内部不正の抑止力となる．さらに，重要な情報資産である名簿を利活用可能になる．

今後の課題として，さらなる詳細な実験行い，2 次元に冗長ビットを付加する手法や，各パラメータの決定手法を研究する必要がある．また，効率的な冗長ビットの生成手法や符号語の 0 のみを冗長に埋め込む手法の検討があげられる．名簿がダウンロードされる際，属性（性別や年齢，住所等）を指定される場合も考慮する必要がある．分割攻撃に関しても，単純に分割するだけでなく属性に偏った分割を行われる可能性もある．これらに対しては，ダウンロードされる名簿の特徴によって，ダミー個人情報を動的に生成する手法も必要になると考えられる．以上が今後の課題である．

参考文献

[1] Symantec Corporation: *2014 Internet Security Threat Report*, Vol.19, Symantec Corporation (2014).

[2] Ponemon Institute LCC: 2014 Global Report on the Cost of Cyber Crime, Ponemon Institute LCC (2014).

[3] 木下　盾，上原哲太郎：追跡可能性を持つ名簿システムにおける結託耐性と分割耐性，コンピュータセキュリティシンポジウム 2016 予稿集 (2016).

[4] 消費者庁消費者制度課：名簿販売事業者における個人情報の提供等に関する実態調査報告書，消費者庁 (2016).

[5] Cappelli, D., Moore, A. and Trzeciak, R.: *The CERT Guide to Insider Threats: How to Prevent, Detect, and Respond to Information Technology Crimes* (*Theft, Sabotage, Fraud*), Addison-Wesley Professional (2012).

[6] 独立行政法人情報処理推進機構技術本部セキュリティセンター：「内部不正による情報セキュリティインシデント実態調査」報告書，独立行政法人情報処理推進機構 (2016).

[7] 独立行政法人情報処理推進機構技術本部セキュリティセンター：組織における内部不正防止ガイドライン，独立行政法人情報処理推進機構 (2017).

[8] 田中賢一，小松尚久：電子透かし技術—ディジタルコンテンツのセキュリティ，p.170，東京電機大学出版局 (2004).

[9] Boneh, D. and Shaw, J.: Collusion-Secure Fingerprinting for Digital Data, *IEEE Trans. Information Theory*, Vol.44, No.5 (1998).

[10] Safavi-Naini, R. and Wang, Y.: Collusion secure q-ary fingerprinting for perceptual content, *Security and Privacy in Digital Rights Management* (*SPDRM'01*), Lecture Notes in Computer Science, Vol.2320, pp.57–75 (2002).

[11] Hollmann, H.D.L., van Lint, J.H., Linnartz, J-P. and Tolhuizen, L.M.G.M.: On codes with the identifiable parent property, *Journalof Combinatorial Theory*, Vol.82, No.2, pp.121–133 (1998).

[12] Staddon, J.N., Stinson, D.R. and Wei, R.: Combinatorial properties of frameproof and traceability codes, *IEEE Trans. Information Theory*, Vol.47, pp.1042–1049 (2001).

[13] Tardos, G.: Optimal probabilistic fingerprint codes, *Proc. 35th Annual ACM Symposium on Theory of Computing*, pp.116–125 (2003).

[14] Trappe, W., Wu, M., Jane Wang, Z. and Ray Liu, K.J.: Anti-collusion Fingerprinting for Multimedia, *IEEE Trans. Signal Processing*, Vol.51, No.4 (2003).

[15] He, S. and Wu, M.: Performance study on multimedia fingerprinting employing traceability codes, LNCS, Vol.3710, pp.84–96 (2005).

[16] Hou, S., Uehara, T., Morimura, Y. and Monoh, M.: Fingerprinting Codes for Live Pay-Television Broadcast Via Internet, *Multimedia Content Analysis and Mining*, pp.252–261 (2007).

[17] Guth, H. and Pfitzmann, B.: Error and collusion-securefingerprinting for digital data, *Information Hiding '99*, Lecture Notes in Computer Science, Vol.1768, pp.134–145, Springer (2000).

[18] Yoshioka, K. and Matsumoto, T.: Random-error-resilient tracing algorithm for a collusion-secure fingerprinting code, *IPSJ Journal*, Vol.43, No.8, pp.2502–2510 (2002).

[19] 今井正樹，上原哲太郎，侯　書会，上田　浩，津田　侑，喜多　一：情報漏洩元の特定を可能とする電子文書管理システム，暗号と情報セキュリティシンポジウム (2012).

[20] 山口哲也：個人情報漏洩監視システム，個人情報漏洩監視方法，個人情報漏洩監視プログラムおよびそのプログラムを記録した記録媒体，特開 2006-79233 (2006).

[21] 根本和郎：データベースアクセス監視装置，情報流出元特定システム，データベースアクセス監視方法，情報流出元特定方法，およびプログラム，特開 2005-222135 (2005).

[22] 佐々木宏：個人情報漏洩監視システム，個人情報漏洩監視方法及び個人情報漏洩監視プログラム，特開 2016-099722 (2016).

[23] 株式会社エターナルコミュニケーションズ：個人情報漏

洩監視システム iTracker，入手先 ⟨http://eternal-communications.co.jp/itracker.html⟩（参照 2016-12-03）．
[24] 今井秀樹：情報理論，p.71，株式会社オーム社 (2014).
[25] Key, J.D.: Some applications of Magma in designs and codes: Oval designs, Hermitian unital and generalized Reed-Muller codes, *J. Symbolic Computation*, pp.37–53 (2001).

木下 盾

2016 年立命館大学情報理工学部情報システム学科卒業．2018 年同大学大学院情報理工学研究科情報理工学専攻博士課程前期課程修了．2018 年より株式会社 NTT データに勤務．

上原 哲太郎 （正会員）

1995 年京都大学大学院工学研究科博士後期課程研究指導認定退学．同大学院工学研究科助手，和歌山大学システム情報学センター講師，京都大学大学院工学研究科助教授，同大学学術情報メディアセンター助教授，総務省情報通信国際戦略局通信規格課標準化推進官を経て，2013 年より立命館大学情報理工学部教授．京都大学博士（工学）．デジタル・フォレンジック，システムセキュリティ等の研究に従事．共著に『IT-Text ネットワークセキュリティ』（オーム社）等．