

偽造防止機能を有するデジタル証明文書システム

A Forgery-Protective Certificate Document System

野崎 剛一*, 河口 英二†, Richard Eason††
Koichi NOZAKI*, Eiji KAWAGUCHI†, and Richard EASON

長崎大学*
Nagasaki University

九州工業大学†
Kyushu Institute of Technology

メイン大学††
University of Maine

情報化が進む今日、文書はデジタル化され、手書きや印刷文書に代わってデジタル文書がインターネットで送受信できるようになった。ところが、今でもデジタル化できない重要な文書がある。それは「証明文書」である。デジタルデータが証明文書になり得ない理由は、データが簡単に改変・偽造され、しかもその証拠が残らないからである。現在、既に「デジタル署名システム」と呼ばれるものがあるが、これは、従来からの公開鍵暗号と電子署名の枠組みを利用するものであり、鍵の取得には「認証局」が必要であって、デジタル文書そのものに偽造防止機能はない。本稿では、ステガノグラフィ技術を使った偽造防止機能を有するデジタル証明文書システムについて述べる。このシステムは、認証局を必要とせずコストのかからないものであり、色々な用途が考えられる。大学や高等学校などにおいては、成績証明書をデジタル文書として発行し、メールで配布することができ、これを誰も改ざんすることができない。

キーワード：ステガノグラフィ、デジタル文書、証明文書、偽造防止、認証局

Almost all documents nowadays can be handled in a digital manner except for one very important case. It is a certificate document. In today's digital age, an authentic certificate document, that is original, cannot take a digital form. The reason why a digital document is not accepted as a certificate is obvious. All digital documents currently used are easily forged without leaving any clues for forgery detection. private key system. However, our new model does not need any authentication bureau. We have already developed the BPCS-Steganography program which is a new steganographic technique using a true color image file (24 bits BMP file) for its information hiding dummy image. In this paper, we propose a new unforgeable digital document system using BPCS-Steganography technique for the document data structure. It is a quite easy-to-use system with very cheap cost.

Keywords : Steganography, digital document, certification, forgery

1. はじめに

一般に、「証明書」とは、社会的に信用度の高い政府機関、国際機関、公共機関、団体、会社法人、有名な個人などが発行したものであり、発行元

の認証作業は必要とされないものである。言い換えれば、そのような信頼度の高い発行元の証明文書が偽造の標的となる。美術品、宝石、骨董品の鑑定書等もこの種の証明書である。

本システムは「デジタル証明文書データ」自身に偽造防止機能を持たせるものである。このシステムでは、例えば成績証明書をデジタル文書で発行し、暗号化せずにインターネットで送付できるし、偽造はすぐに発見される。この証明書は実際には「画像ファイル」であり、いくらでもコピーが可能で、紛失の恐れもない。

その原理は非常に単純である。即ち、現在のデジタル文書データは、外部から見える（読み取れる）情報データだけの「外部層」しか持っていない。したがって、それを改変、差し替え、贋造することは容易である。これを防ぐにはデジタル文書データの情報構造を「多層化」すればよい。多層化してその内部層が外部から見えないようにして、内部に外部層と同じ情報を埋め込み内部層

*総合情報処理センター

〒852-8521 長崎市文教町1番14号
Information Science Center, Nagasaki University
〒852-8521 1-14 Bunkyo-machi, Nagasaki
E-mail : nozaki@net.nagasaki-u.ac.jp

†工学部電気工学科

〒804-8550 福岡県北九州市戸畑区仙水町1-1
Kyushu Institute of Technology
〒804-8550 1-1 Sensui-cho, Tobata-ku, Kitakyushu, JAPAN
E-mail : kawaguch@know.comp.kyutech.ac.jp

††Dept. of Electrical and Computer Engineering,
University of Maine, 5708 Barrows Hall, Orono, Maine 04473-5708, U.S.A.

を改変できないようにすればよい。そして、真正性の判定には、内部層の情報を取り出して外部層と比較すればよい。外部と内部の情報一致すれば改ざんのない真正文書であり、違っていたら外部を改ざんした文書である。文書情報の多層化を可能にする技術としては、著者等が開発したBPCS-Steganography 技術^{1), 2)} が最適である。

2. BPCS-Steganography技術

2.1 ステガノグラフィ

既にインターネットは日常生活に欠かせないものとなり、情報セキュリティに関する社会の関心も徐々に高くなりつつある。最近、セキュリティ技術の一つとしてステガノグラフィが注目され始めたが、まだ一般にはなじみが薄い。

ステガノグラフィ (Steganography) とは、他人に知られること無く情報を伝達することを意味する言葉である。従来、ステガノグラフィは物理化学的な現象を利用したアナログ技術であったが、コンピュータによるマルチメディアデータ処理が可能となった今日では、デジタルステガノグラフィを意味し、「インフォメーション・ハイディング技術」の一部になっている。

2.2 暗号や電子透かしとの違い

一般に、コンピュータファイルを保護する方法としては暗号技術がある。情報の暗号化とは、秘密情報を「読めなくする」技術であり、既にインターネット上でも実用化されている。しかし暗号は、そこに「秘密データがある」ことを隠すものではない。一方、ステガノグラフィとは、暗号化されている・いないにかかわらず、「情報の存在」そのものを他人の目に「見えないように」するものである。即ち、秘密の存在を「非可視化」する技術であると言える。

ところで、このステガノグラフィ技術と似た Watermarking, 画像深層暗号化などと呼ばれる「電子透かし技術」は、電子データに関するオリジナリティや所有権の主張あるいは認証のために、少量の「目印情報」をその電子データに埋め込んでおく技術である。この場合の「価値有る情報」とは、表に現われている電子データそのものである。これに対して、ステガノグラフィでは、表に現われるデータはあまり価値がないダミーのデータであり、中に埋め込まれたものに価値がある。

このように、電子透かしとステガノグラフィは、用途と目的が相反するインフォメーション・ハイディング技術である。ステガノグラフィと電子透かしの違いおよび位置づけをまとめると、表1と図1のようになる³⁾。

表1. ステガノグラフィと電子透かしの違い

	ステガノグラフィ	電子透かし
価値ある情報	埋め込まれた情報	外に表れた情報
埋め込みデータの頑強さ	外に表れた情報を加工すると、内部に隠された情報は壊れても構わない	外に表れた情報を加工しても取り除けない
埋め込み容量	できるだけ大きいことが望ましい	少容量でよい

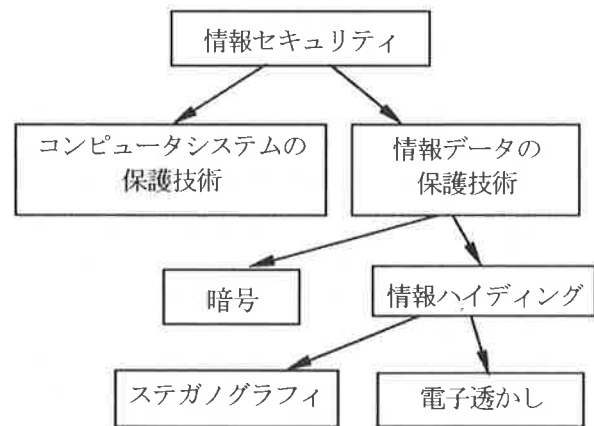


図1 ステガノグラフィ技術の位置づけ

2.3 BPCS-Steganography

BPCS-Steganography とは、画像データを各構成ビットにスライスして得られるビットプレーンに関して、そのプレーン上の2値パターンの複雑さ³⁾に注目して、複雑な部分に秘密データを埋め込む(データを置き換える)技術 (Bit Plane Complexity Segmentation Steganography) として著者等が開発した方法である。

2.3.1 2値画像データの複雑さの定義

0と1から成る2値画像データを2次元配列のデータとして捕らえて、0と1の変化を縦方向と横方向で足し合わせた数値を「2値画像の境界線の長さ」とする。例えば、白画素で囲まれた孤立した1個の黒画素の境界線長は4であり、8×8

画素からなる次の図2に示す白黒2値パターンPの境界線の長さは24で、市松模様の場合では112である。

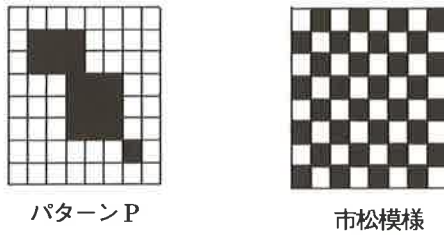


図2 白黒2値パターンの例と市松模様

今、 $m \times m$ 画素からなる2値画像を考え、その画像内において白黒の色の変わり目を縦方向と横方向について数える。境界線の最小値は0である(全て白、又は全て黒の場合)。一方、境界線の最大値は、 $2m(m-1)$ で、市松模様の場合に得られる。この境界線の長さを利用して、「複雑さ α 」を以下の式で定義する。

$$\alpha = \frac{k}{2m(m-1)}$$

k : 2値画像内の境界線の長さの値

2.3.2 基本原理

自然画像のビットプレーンでは、ノイズ状のデータを別のノイズ状のデータで置換しても視覚的にはほとんど影響がない。特に、24ビットフルカラーのBMP形式画像ファイルデータの場合、RGBの第5ビット目以降のビットプレーンについては、殆どがノイズ状の領域である(図3参照)ので、この領域を秘密データで置換することができる。ノイズ状の領域であるか否かの判断基準として、複雑さの尺度 α を用いる。この値は、画像によって異なり、それぞれに適した閾値を定める必要がある。

2値画像において、例えば 8×8 画素からなる小領域について、閾値 α_{TH} に対して、

$$\alpha_{TH} \leq \alpha$$

を満たす領域を秘密データの埋め込み場所とする。

2値画像のノイズ状の領域に秘密データを埋め込むには、まず、その秘密データをランダム化し(圧縮すればランダムになる)、 8×8 ビット毎に区切り、ファイルを小画像の系列とみなして、それらの小画像を順次、ダミー画像上の 8×8 のノイズ状の領域に埋め込んでいく。埋め込む小画像がノイズ状でない場合にはコンジュゲーション演算により複雑化(ノイズ状のデータへ変換)

する。このような操作により、どのような秘密データもダミー画像に埋め込むことが可能となる。ただし、この場合、埋め込んだファイルを完全に復元するために、コンジュゲーション演算を施した領域の位置の記録「コンジュゲーションマップ」を保存しておかなければならない。



24ビットカラー画像

24ビットカラー画像のRed成分(8ビット)のビットプレーン分解画像を以下に示す。下の図の左側上より第1, 2, 3, 4, そして右側上より第5, 6, 7, 8ビットプレーン画像である。

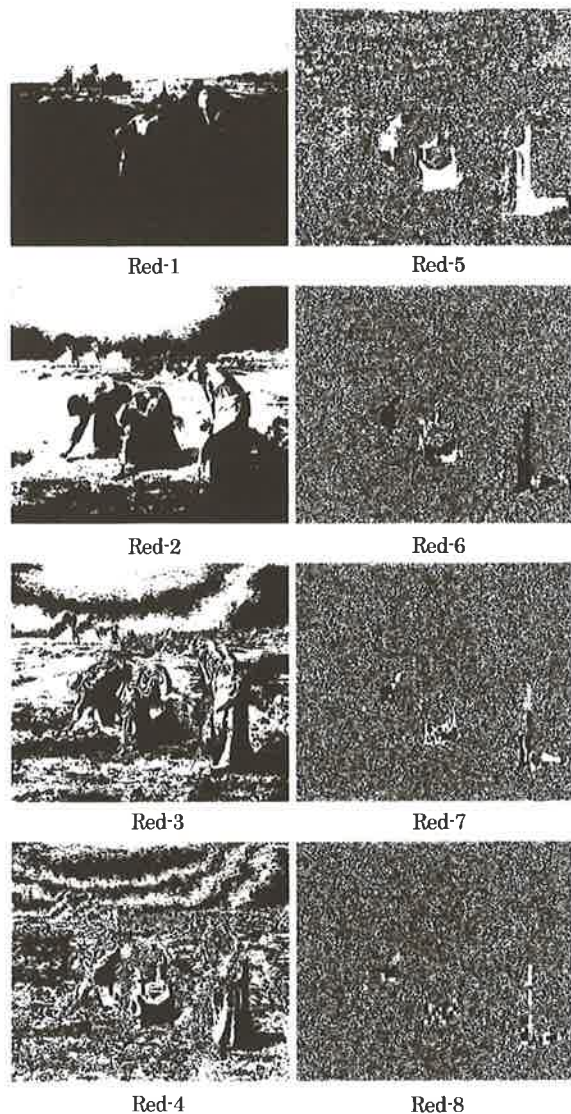


図3 画像ファイルのRGB8ビット分解例

2.4 BPCS-Steganography の特徴

BPCS-Steganography 技術は複雑さの尺度とコンジュゲーション操作など独自のアルゴリズムを有する⁴⁾。次に、本システムに関してその特徴を示す。

(1) 大量情報の埋め込み

ダミーデータを 24 ビットフルカラーの BMP 形式画像ファイルとする場合、ダミー画像データファイルの容量の約 50%までデータ埋め込みが可能である。これは、Web 上に公開されている他のステガノグラフィプログラムのデータ埋め込み容量の 2 倍以上の埋め込み能力を持つ⁵⁾。

(2) 壊れ易い

このプログラムで取り扱うデータは外見上単なる画像データである。外に見えている画像データを画像処理ソフトで加工、改ざんすると、内部のデータ構造が変化するので、埋め込まれたデータを取り出すことができなくなる。即ち、画像データの改ざんを検出できる。

(3) 埋め込まれたデータは外から改変できない

データは外層と内層の二重構造となっていて、データ埋め込みプログラムの使用者のみが持つ埋め込み鍵とデータ埋め込みプログラムがない限りそのデータの生成や複製はできない。

3. デジタル証明文書システム

3.1 デジタル証明文書

「証明書」はオリジナルが絶対に改変されてはいけない文書である。コンピュータで扱う「デジタル文書データ」は、テキスト、記号、表、図形、イラスト図、写真等から成る。そしてこれら

が統合されてディスプレイに表示される時には、紙の文書のような見える（読める）形の情報となる。このようなデジタル文書は、編集可能なファイル（Word ファイル、画像ファイル等）か、部分的に編集可能なファイル（PDF ファイル等）である。従って、オリジナル文書を後で改変することは簡単であり、通常のデジタル文書そのまま「証明文書」にすることはできない。

3.2 偽造文書の分類

紙などの証明文書の偽造は、以下の 2 つのタイプに分けられる。

(1) タイプ 1：真正な文書の一部を改変するもの。

たとえば、写真付きのクレジットカードの写真を張り替えて他人に成りすますことや大学の学業成績証明書の一部を改ざんして優秀な成績に見せかけること。精巧な偽造は見抜けない。

(2) タイプ 2：文書全体の情報内容を偽造するもの。

たとえば、パスポートの偽造では、偽造台紙に、不正なデータを精巧に記載すること。真正か偽造かは、台紙の質や書式だけでしか判断できない。

3.3 電子署名・認証の仕組み

デジタル画像データの加工・編集は、市販のソフトを使えば簡単にできる。そのためデジタル画像データの原本性は、従来の銀塩写真や紙の書類に比較して低く、証拠としての能力に乏しい。

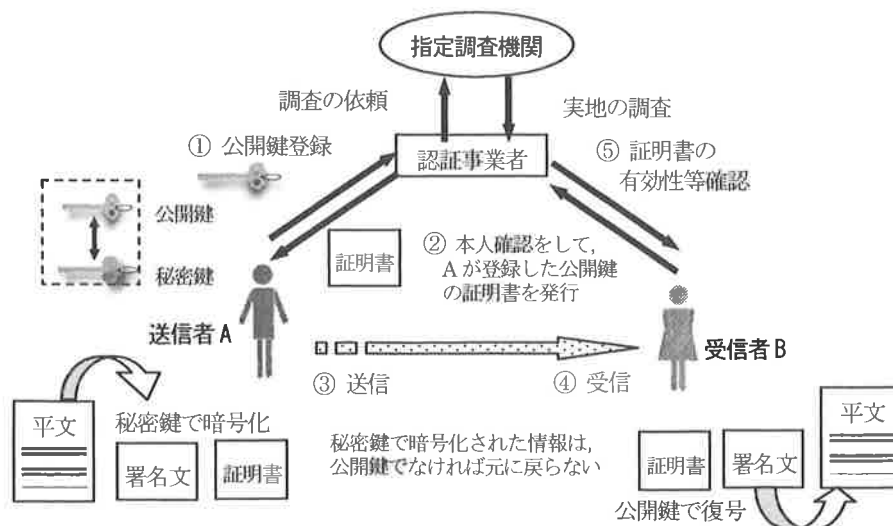


図4 電子署名・認証の仕組み

そのために、デジタル署名（電子署名）で使う公開鍵を入手し、デジタル証明書を利用することで相手を正しく識別する方法が用いられる。例えば、手書き署名や押印と同等に通用する法的基盤を整備するために、日本では電子署名法が2001年4月1日から施行された。また、認証業務のうち一定の基準を満たすものは、国の認定を受けることができる制度が導入された。

その電子署名・認証の仕組み（図4参照）⁶⁾には、公開鍵方式が利用されている。この方式ではデジタル文書そのものには認証力（即ち、真正であることの証明能力）はない。このデジタル署名のPKI（Public Key Infrastructure）「公開鍵暗号基盤」は、日本の印鑑登録制度と非常に似ているが、両者は、まったく別の制度である。印鑑登録は慣習に基礎を置く制度で、日本では不動産や自動車の登録、購入の取引において利用されているが、一般の消費者取引においては使われない。一方、インターネットでの消費者取引においては、印鑑証明に匹敵するようなセキュリティが求められ、PKIが利用されているが、公的な認証局を必要とするために使い難い。今後、デジタル文書をコンピュータネットワークでやり取りすることは盛んになり、証明文書のデジタル化の必要性が高くなっていくことは確実である。更に使い易い仕組みと偽造防止機能が必要となる。

3.4 デジタル文書データの多層化情報構造

デジタルステガノグラフィ技術によりデータが埋め込まれた画像ファイルは、改ざん編集により内部のデータが壊れやすい特徴を持つ。この特徴を利用すれば、文書を画像データ化することによって文書の原本性を確認することが可能になる。

次に、これを実現する原理について述べる。

デジタル文書は、コンピュータ・ディスプレイで見える（読める）ように作成されており、人の目には、どのようなデジタル文書も「形や色

を持った画像」として映るので、文書はすべて「画像データ」であっても構わない。デジタルステガノグラフィ技術で埋め込まれたデータ（内部層の情報）は外部からは全く見えないし、外部に見える画像（外部層）も劣化しない。外部層を操作して、埋め込まれたデータを書き換えることは不可能である。無理に外部を改変すると、必ず、内部層とは違ったものになる。

多層化情報構造を持つデジタル文書データとは、見える（読める）ままのオリジナル文書情報を外部層に持ち、同時に内部層にはオリジナル文書と埋め込み鍵（Key_p）を攪拌した鍵（公開抽出鍵：Key_v）が埋め込まれているものである。そのような多層化データ構造はステガノグラフィ技術によって実現できるし、検証のために後で内部情報を取り出すこともできる（図5参照）。

3.5 真正性の判定

多層化情報構造を持つデジタル文書が改ざんされていない（真正な文書）とは、外部層と内部層の情報が完全一致する（または内部層が外部層と照合している）ことであり、そうでなければ改ざんされた文書である。本システムの多層化構造のデジタル文書を受け取った者は、その文書の真正性を確認するために、オリジナル文書の発行元から公開抽出鍵（Key_v）を受け取り、内部層のデータ抽出処理を行う。鍵の受け取りは、Webページやメールなどにより行う。抽出プログラムにより改ざんが発見されると不正なデータであることになる。抽出した内部層データは、外部層データと目視により比較して真正性の確認を行うことができる。また、外部層データと内部層データのイメージデータをOCRソフトで文字認識させて、両者の比較で改ざん箇所を発見することも可能である。このデジタル証明文書の作成と真正性の判定の流れを図6に示す。

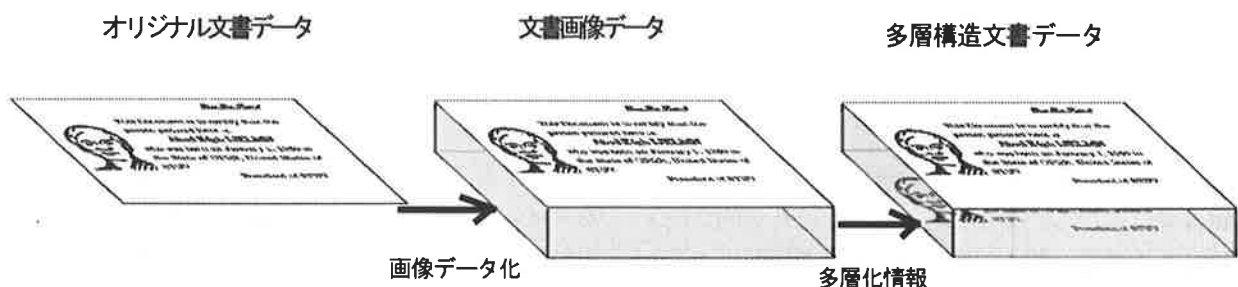


図5 多層構造文書データ

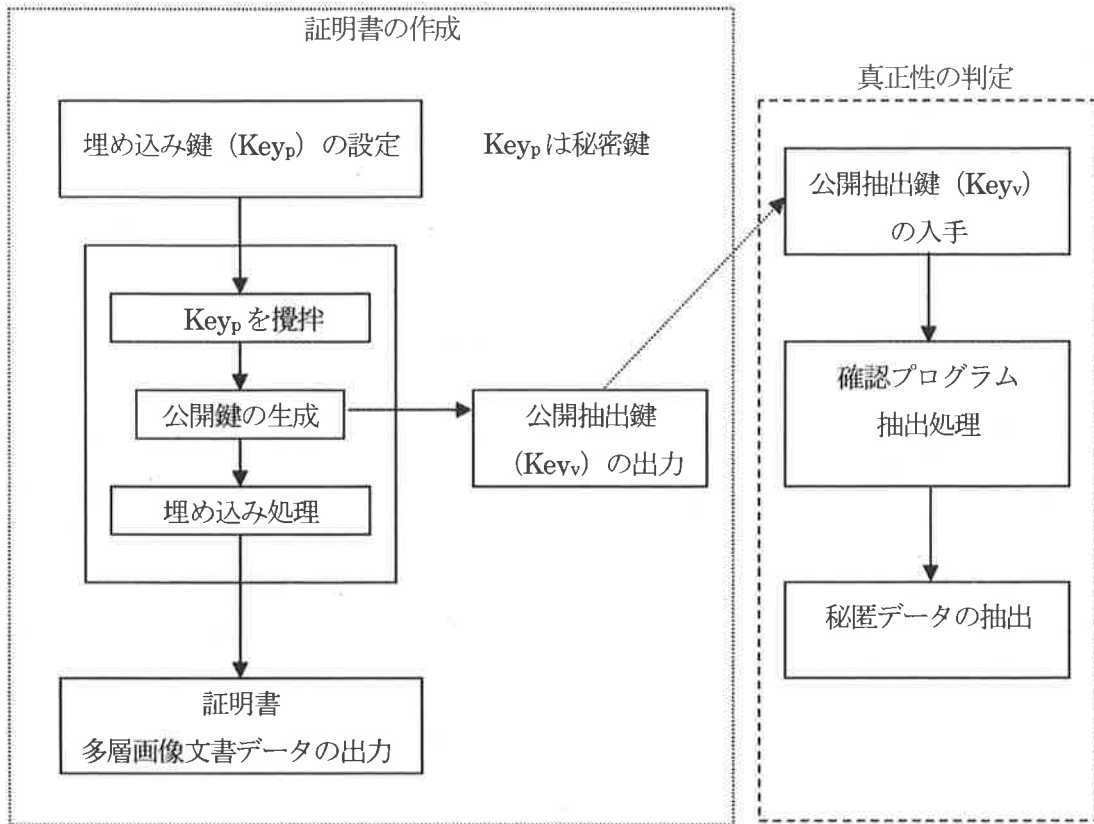


図6 デジタル証明書の仕組み

4. 3種の基本システム

偽造防止機能付きデジタル証明文書システムは、用途毎に以下の3つのタイプに分かれる。

- (1) 1型システム:タイプ-1の偽造を防止するシステム

これは文書が発行（作成）された後で、何らかの変更が行われたかどうかだけを判定するシステムである。最も基本的なシステムであり、証明文書作成サブシステムと検証サブシステムが対となり、鍵は使用しない。検証サブシステムは一般に公開する。このシステムはデジタル・クレジットカードなどに適用できる。

- (2) 2型システム:タイプ-2の偽造防止機能を持つシステム

1型システムに、証明文書作成鍵と検証鍵を設定したものである。証明文書は作成鍵を使って作成され、検証は（即ち、証明文書の内部層を抽出することは）検証鍵を使って行う。作成鍵は証明文書作成元（発行元）に秘密裏に保管され、検証鍵は発行元から公開される。発行元は社会的に信頼のある組織・個人であるので第三者機関による認証は必要とされない。検証鍵は、文書作成サブ

システム中で、外部に見えないようにして作成鍵から生成されるが、これは一方向スクランブリング関数（例えば、入力値によってスクランブルのアルゴリズムを変える方法）で実現される。従って、検証鍵から作成鍵を探し出すことは「現実的に不可能」である。このような文書を偽造するには、正規の発行元から公開された検証鍵で検証できるように、正規の発行元だけしか知り得ない作成鍵を使う必要がある。しかしその鍵は秘密であるので偽造はできない。このシステムはデジタルパスポートなどに適用できる。

- (3) 3型システム:オリジナル文書の作成者を判定できるシステム

2型システムの文書生成時に生成時刻情報（インターネット・タイムスタンプ）を付加したシステムであり、タイプ-1、2の偽造防止機能に加え、オリジナル作成者（誰が最初にその文書を作成したか）が判定できる。タイムスタンプ自身も外部層、内部層情報として扱うので、最も早い作成者（その作成鍵の所有者）が判定できる。外部層のタイムスタンプを改ざんしても内部層のタイムスタンプを改ざんすることは不可能である。このシステムは、撮影した日時と撮影画像内容の改ざんを防止できるデジタル写真に適用できる。

これによりデジタルな証拠写真が実現される。

5. デジタル証明文書システムの特徴と応用

本システムの特徴は、以下の通りである。

- 証明文書作成サブシステムと検証サブシステムのみで動作するので、利用方法が簡易であること
- デジタル文書（画像ファイル）の変造、偽造が不可能であること（偽造すると、埋め込まれたデータが壊れる）
- 認証局のような仕組みを使わなくても利用が可能なこと
- 埋め込み鍵と抽出鍵を分離できる（埋め込み鍵：非公開、抽出鍵：公開）
- オンサイト、オフラインで判定できること
- 公の認証局が要らないので、あまりコストがかからないこと

具体的な応用が色々と考えられるが、次に2つの事例を示す。まず第一番目の例として、大学における成績証明書をデジタル多層化文書として発行するような利用が考えられる。デジタル成績証明書を発行する大学は、成績証明書の文書ファイルを画像ファイルに変換し、多層構造文書を作成する。この文書生成には、大学固有の文書生成鍵を使用して、公開抽出鍵を生成しておく。学生の就職試験において、成績証明書をデジタル多層化文書ファイルとして受け取った企業等は、データ抽出確認プログラムと証明書発行元の大学から公開抽出鍵を入手すれば、文書の真正性の確認を行うことができる。

次に、第二番目の例として、ICメモリ付きデジタルパスポートの例（図7参照）を示す。これは、現在のパスポートの1ページをプラスチックカード表面に印刷してICメモリを付けた形式にした例である。このICメモリ内には、カード表面に印刷された顔写真（IDカードの所有者本人の顔写真）などの情報を多重化情報構造のデジタルデータとして格納しておく。このIDカードが本物であるかどうかのチェックは、このICメモリ読み取り装置とデータ抽出確認プログラムを使って行う。即ち、まず、ICメモリ内の多重化情報構造のデジタル画像データを取り出す。この画像データがIDカードの表に見える顔写真を含む情報と同一でないならば、このIDカードは

本物ではない。この多層構造の画像ファイルから抽出した内部層のデータがIDカードの表の情報と一致する場合に限り、本物である。ICメモリの内容改ざんが行われた場合には、内部層データは壊れてしまうので、改ざんを検出できる。このように、ICカードメモリ内の多層構造画像ファイルの内部層データとIDカード表面のデータが一致することを検証することで本物であるか否かを判定できる。このデータの検証には、ICメモリ読み取り装置とデータ抽出確認プログラムがあればどこでも利用できる。

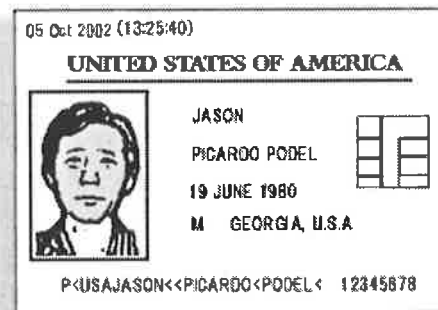


図7 ICメモリ付きデジタルパスポートの例

6. おわりに

最近、PKIと電子署名により、文書の真正性の確認が行われ始めたが、一般に使い難いためになかなか普及していない。また、公開鍵方式を利用する場合は、受信者の公開鍵の取得と認証のために、送信者はまず認証局との対応が必要である。この場合も、認証局が正常に運営されていたとしても、利用の手続き等に時間と労力を要する。しかも、実際には世界的な規模で運用されている認証局はまだ存在しないのが現状である。本システムは、これまでのような認証局を必要としないで、文書の偽造防止機能を持つ新たなもので、ネットワーク社会におけるデジタル証明文書のニーズに対応できる仕組みである。クレジットカードに適用すれば、カードの偽造を発見することができるなど、色々な活用方法が考えられる。

ただし、本システムが機能するためには、データ埋め込みプログラムと抽出プログラムが解読され、作り変えられないことが絶対条件であるので、本システムのプログラムについては、そのアルゴリズムを公開しない。

今日、リバースエンジニアリングの研究により、実行形式のプログラムファイルからある程度のソースコードの復元が可能であり、そのためのリバースエンジニアリング・ツールが世の中に流通している。このような状況の中で応用システムの安全性を確保するためには、処理の詳細がプログラムの作成者以外には理解し難い方法(難解な方法)でプログラムを作成することが必要となる。このことは、一般にコンピュータプログラムについて、プログラム作成者以外の者にも処理の内容が理解し易いように整然と作ることが求められていることとは、正反対のことを行うことを意味する。解析が困難なプログラムの作成技術は、いくつか提案されていて、「プログラムの難読化」、「プログラムの暗号化」、「プログラムの断片化」の3つに大別できる⁷⁾。いずれの方法もプログラムの解析にかかるコスト(時間と労力)を増大させる効果があり、解析を著しく困難にすることが可能である。即ち、プログラムの解析コストが十分に高ければ、そのプログラムの解析に対する意欲を失わせることができる。

このように、解読されたくないソフトウェアやアルゴリズムを公開しないソフトウェアについては、プログラムの解析や改ざん等の行為を困難にする工夫が必要となるが、本システムのプログラムを難解にする具体的な方法については、これを公表することはできない。

参考文献

- (1) Koichi Nozaki, et al : "A Large Capacity Steganography Using Color BMP Images", Proc. ACCV'98, pp.112-119 (1998-01)
- (2) URL : <http://www.know.comp.kyutech.ac.jp/BPCSe/>
- (3) 河口英二, 野田秀樹, 新見道治: "画像を用いたステガノグラフィ", 情報処理 44 卷 3 号, pp.236-241 (2003-03)
- (4) Koichi Nozaki, Michiharu Niimi, Richard O.Eason, Eiji Kawaguchi : "A Large Capacity Steganography Using Color BMP Images", Lecture Notes in Computer Science 1351, Computer Vision ACCV'98 pp.112-119(1998-01)
- (5) 野崎剛一, 新見道治, 野田秀樹, 河口英二 : "BMP ファイルを使った Steganography", 大学情報システム環境研究, Vol.1, pp.16-23 (1998-04)
- (6) 法務省民事局 URL : <http://www.moj.go.jp/MINJI/minji32-1.html>, "電子署名法の概要について", (2003-11)
- (7) 神崎雄一郎, 門田暁人, 中村匡秀, 松本健一 : "命令コードの実行時置き換えによるプログラムの解析防止", 電子情報通信学会技術報告, 情報セキュリティ研究会, Vol. ISEC 2002-98, pp.13-19 (2002-12)

(2003年11月14日原稿受付)

(2004年3月4日採録決定)

著者略歴



野崎 剛一

1975年九州大学工学部卒業, 1975年長崎大学助手, 1982年長崎大学情報処理センター講師, 現在, 長崎大学総合情報処理センター助教授, プログラ

ミングツール, 教育研究用計算機ネットワークシステム, 自然言語概念の意味処理, 情報処理教育, ステガノグラフィ, 画像処理システムに関する研究に従事。

河口 英二

1969年九州大学大学院博士課程単位取得退学, 1969年九州産業大学講師, 1971年九州大学工学部助教授, 1988年九州工業大学工学部電気工学科教授, 言語処理, パターン認識, 画像処理, ステガノグラフィ, 自然言語の意味表現の研究に従事。工学博士

Richard Eason

1978年米国テネシー州立大学工学部電気工学科卒業, 1978年リサーチアシスタント, 1980年同大学同学科修士課程修了, 1980年Zilog社勤務 VLSI エンジニア, 1982年米国テネシー州立大学リサーチアシスタント, 1988年米国テネシー州立大学電気工学 Ph.D., 1988年米国メイン州立大学計算機工学科準教授, VLSI デザイン, マイクロプロセッサ応用, 人工知能, パターン認識, ニューラルネットワーク, ステガノグラフィの研究に従事。