

## テキスト秘密分散 Text Secret Sharing

山村明弘\*  
Akihiro Yamamura

滝澤修\*  
Osamu Takizawa

**あらまし** 自然言語文に秘密を分散させることにより、分散データの存在そのものを隠してしまう秘密分散法について考察する。秘密データの復号方法には自然言語が本来持っているセマンティクスを閾値に利用する。テキストへの秘密分散は視覚復号型秘密分散とおなじような考えをもっている一方で、分散データが自然言語へ埋め込まれるため、それがステガノグラフィとしての役割を持つことを可能にしている点が既存の秘密分散と異なる点である。また日本語の形態素解析ソフトウェアを利用した実験について報告する。またテキスト秘密分散の数学的なモデル化とアクセス構造の構成法について考察する。

**キーワード** 秘密分散、視覚復号型秘密分散、自然言語処理、形態素解析、ステガノグラフィ

### 1 はじめに

複数のメンバーが分散して保有する情報を合わせた場合にのみ秘密情報を復号できる秘密分散法 (secret sharing) <sup>[1][2]</sup> の一つの実現形態として、Naorら <sup>[3]</sup> によって提案された視覚復号型秘密分散法 (Visual Cryptography) は、計算機を使わず人間の視認によって復号可能な新しい暗号として、研究や実用化が進められている <sup>[4][5][6]</sup>。

ところで有史以来の古典的な暗号は、自然言語テキストを対象とするものがほとんどを占めていた。画像におけるピクセルに対応するのは、自然言語テキストの場合は文字である。しかし自然言語テキストを情報隠蔽媒体とするには、文字コードには冗長性が全く無く <sup>[7]</sup>、また言葉は意味をもつことから、情報を隠蔽するために文字を改変すれば僅かな改変でも露見してしまう難点がある。そのため、画像におけるピクセル単位の操作のような方法は文字には単純には適用できない。また自然言語テキストは、画像等の他の媒体と比べて情報量が少ないため、隠蔽できる情報量も少なく、実用的に不十分な場合が多い。そのため、自然言語テキストを情報ハイディングや秘密分散などの情報隠蔽媒体とする研究は、一部の例外 <sup>[8][9][10]</sup> を除き、あまり多く見られない。しかし、マルチメディア化が進んでいる現代においても電子メールなど自然言語テキストでの情報交換は主流の位置を占めており、情報伝達手段としての自然言語テキストの重要性は今後も変わらないと考えられる。従って自然言語テキストを情報隠蔽媒体として扱う暗号法には、多くの応用が期待できる。そこで本稿では、自然言語テキストを情報隠蔽媒体とする秘密分散法 (Text Secret Sharing) の実現の可能性について検討し、一つの

簡単な方法を提案する。

### 2 テキスト秘密分散の考え方

視覚復号型秘密分散の考え方を自然言語文に適用した新しい秘密分散法を提案する。日本語文を対象とし、複数枚の分散テキスト (share text) を重ね合わせて、1文字目を上層から下層、2文字目を上層から下層、... に順次並べていくと、その文字列の中に秘密テキスト (secret text) が現れるようにするものである。積み重ねる順序を変えることによって、別の秘密テキストを浮かび上がらせることも可能である。分散テキストを重ね合わせることは簡単な機械処理によって実現できる。重ね合わせて得られた文字列の中から秘密テキストを復号することは、人間による視認でも可能だが、意味のあるフレーズは2文字以上の形態素の連なりになっている確率が高い性質を利用して、形態素解析器を援用して切り出す方法を提案する。分散テキストは、文データベースを用いて合成することで、一見自然な文に見せかけることができる。

テキスト秘密分散法は、秘密テキスト (secret text) を複数の分散テキスト (share text) に分散して隠蔽し、テキストを“重ね合わせて”秘密テキストを復号するものと定義できる。視覚復号型秘密分散法では、分散画像はノイズのような無意味画像が使われることが多い <sup>[11]</sup>。従ってテキスト秘密分散法の場合も分散テキストが無意味な文字列であっても構わないが、そうすると秘密分散を使っていることを見抜かれる懸念があり、それは既に脅威となりえることなので、自然言語処理的な工夫により、できるだけ自然な文にすることが望ましいといえる。

ところで、テキスト秘密分散法において、視覚復号型秘密分散法における“重ね合わせる”ことに対応するのは

\* 独立行政法人 通信総合研究所 情報通信部門 〒184-8795 東京都小金井市貫井北町 4-2-1  
Communications Research Laboratory, Information and Network Systems Division, 4-2-1, Nukui-Kitamachi, Koganei, Tokyo 184-8795 Japan

どのような処理であろうか？ 例として以下のような方法が考えられる。

(方法1)

分散テキストを複数枚重ね合わせて、一致している文字を拾い読みすると、秘密テキストになっているようにする方法が考えられる。図1にイメージを示す。

秘(盛)機(機)筋(湯)に(測)羅(部)の(粒)計(が)原(初)察(察)  
蔵(高)階(階)弁(弁)駅(住)電(電)機(機)修(用)欄(欄)程(議)前(皮)  
波(が)種(種)作(作)を(電)送(送)る(る)が(電)会(会)類(類)業(業)研(研)技(技)術(術)  
電(電)離(離)碑(碑)の(の)数(数)機(機)場(場)編(編)随(随)離(離)電(電)の(の)茨(茨)蒂(蒂)集(集)  
社(社)種(種)層(層)電(電)新(新)城(城)を(を)重(重)重(重)再(再)重(重)信(信)上(上)

図1 方法1のイメージ

分散テキストを重ね合わせて上から眺めた場合 (丸印を付した個所が秘密テキスト「盛岡駅前が会場です」)

この方法は、文字をイメージとして扱っている点で、視覚復号型秘密分散法の一つに位置づけることができ、視覚復号型秘密分散法のように視認で復号できる特長がある。重ね合わせる順序は鍵にならない。また、分散テキストが完全に揃っていない場合でも、秘密テキストは完全ではない状態(過剰な文字が混入している)であるものの、過剰文字を問引いて読むことで大まかに解読できてしまう懸念がある。

(方法2)

複数枚の分散テキストを重ね合わせて、冒頭文字の位置を合わせ、1文字目を上層から下層、2文字目を上層から下層、...に順次並べていくと、その文字列の中に秘密テキストが現れるようにする方法が考えられる。図2にイメージを示す。

分散テキスト1 ...方向の臨界周波数<sup>盛</sup>目の0MHzを示す。...  
分散テキスト2 ...温秋田電波観測所山岡己雄郵政大臣表彰を...  
: ...、国鉄東北本線古河駅より東、筑波山に向...  
...土達により10年程前に紹介されている。...  
...いう面で、その施策が不十分であったと認...  
...、別に関係する各学会の雑誌もあり、また...  
...な研究成果の発表の場としては、別に関係...  
...うとの意に出たものである。最近は、本来...  
...究所ニュース」と題する一般広報用小冊子...  
...行することになった。皆さんは何と聞くで...

図2 方法2のイメージ

分散テキストを重ね合わせて横から眺めた場合(四角で囲った範囲が秘密テキスト「盛岡駅前が会場です」)

重ね合わせて得られる文字列は、無意味な文字列の一部

に、意味のある秘密テキストが混じっている形になる。図2の場合、重ね合わせて得られる文字列は、縦に左から右へ読んでいくことで得られ、「...波測線0のる発た」な数所古年施各表もとっ目山河程策学のの題た盛岡駅前が会場です。の己よに不のとある0雄り紹十雑しる一M郵東介分...」(一部)となる(下線部が秘密テキスト)。この方法は、重ね合わせる順序も鍵になっているので、順序を入れ替えることで別の秘密テキストも隠蔽することが原理的には可能である。但し分散テキストが完全に揃っていない場合でも、秘密テキストは完全ではない状態(一部の文字が歯抜け)であるものの、大まかに解読できてしまう懸念がある。

この方法は、重ね合わせる行為自体に計算機の補助が必要になる。また秘密テキストが隠蔽されている箇所は視認で抽出できなくはないが、結構面倒なため、やはり計算機の補助が必要となる。従って、物理的な重ね合わせ操作および視認によって直ちに復号できる視覚復号型秘密分散法のような簡便さに欠けている難点がある。

以下の節では、方法2の可能性について更に詳しく検討する。

### 3 テキスト秘密分散の一方法

本節では、前節の方法2を実現した簡単な処理プログラムを実装した結果について述べる。

#### 3.1 分散テキストの生成方法

第2節で述べたように、分散テキストは自然な文章であることが望ましい。単語を組み合わせることによって意味的に自然な文を合成することは、計算機では膨大な知識と複雑なアルゴリズムを必要とする。そこで、ある程度の長さの文を大量にデータベース化しておき、その文を組み合わせることによって分散テキストとする方法を考える。ここで使うデータベースは、2つ以上の文を続けても文間の文脈が極端には不自然にならないように、同じ分野の文によって構成されていることが重要と考えられる。

分散テキストを重ねて秘密テキストがうまく現れるようにする方法を考える。例えば、「滝澤より、文京グリーンコートセンターオフィス16階でお待ちしています。」という秘密テキストを10枚の分散テキストに1箇所に隠蔽させることを想定すると、それぞれの分散テキストの中に

- 分散テキスト1 「...滝ンて...
- 分散テキスト2 「...澤コスイ...
- 分散テキスト3 「...よー1ま...
- 分散テキスト4 「...リト6す...
- 分散テキスト5 「...セ階。...
- 分散テキスト6 「...文で...
- 分散テキスト7 「...京タお...
- 分散テキスト8 「...グー待...
- 分散テキスト9 「...リオち...
- 分散テキスト10 「...ーフし...

という、それぞれ3~4文字のフレーズを同じ位置に入れなければならないことになる。例えば分散テキスト1の場合、「滝ンて」という無意味なフレーズを内包させて、どのように分散テキストを作るかが問題となる。以下の対処方

法が考えられる。

(対処1) 分散テキスト中に多少おかしなフレーズがあっても許容する

秘密テキストの正確な復号を実現するために、分散テキストの自然さを犠牲にする方法である。例えば上掲の例における分散テキスト1の「滝ンにて」を内包した文を無理やり合成する。この方法では、おかしなフレーズがある場所を手がかりとして隠蔽情報の存在を見破られる危険性がある。

(対処2) 復号された秘密テキストが原文とは多少異なっても許容する

分散テキストの自然さを優先して、秘密テキストの正確な復号を犠牲にする方法である。例えば上掲の例における分散テキスト1の「滝ンにて」を「滝シにて」程度に調整することを許容する。この方法では、秘密テキストが形態素的におかしな文になるので、3.2項で述べる形態素解析を用いた復号が困難になることが問題となる。またこの調整の自動化は難しいと思われる。

(対処3) 秘密テキストをちぎる

秘密テキストを複数の部分秘密テキストにちぎって、「部分秘密テキストの最大文字数 分散テキストの枚数」とし、分散テキスト上で置く位置を散らせる。但しちぎる際に形態素を分断するような変なちぎり方をすると、3.2項で述べる形態素解析を用いた復号が困難になるので、工夫が必要である。また、置く位置を散らせた状態でうまく分散テキストを合成することは難しい。

(対処4) 秘密テキストの文字数 分散テキストの枚数とする

最も簡単に実現できる方法であるが、十分な長さの秘密テキストを使えないか、もしくは大量の分散テキストを使わなければならない等の点で、実用上の大きな制約がある。

本稿では、とりあえず対処4を採用することにする。

### 3.2 秘密テキストの復号支援方法

第2節で述べた通り、方法2は秘密テキストが隠蔽されている箇所を抽出することが視認では比較的難しいため、対策として例えば秘密テキストの開始と終了の箇所にフラグシーケンスを付加する方法が考えられる。しかし、第1節で述べたように自然言語テキストを情報隠蔽の媒体として用いる場合、隠蔽できる情報量が少ない制約があるため、秘密テキストもできるだけ短いことが望ましい。そこで、意味のあるフレーズとそうでないフレーズとは視認で見分けがある程度可能という自然言語の性質を利用することで、フラグシーケンスを使わなくても抽出できるような手立てを考える。

自然言語処理における基本的な処理の一つとして、文を形態素(語を構成する最小単位)に分解する形態素解析がある。形態素解析をすると、意味の無いフレーズは、1文字の形態素の並びになる場合が圧倒的に多い。そのた

め、2文字以上の形態素が多く現れる箇所は意味のある秘密テキストの部分である可能性が高いことになる。この性質を復号に援用する。

### 3.3 具体的な処理の例

3.1項の対処4に基づく分散テキストの生成機能、および3.2項の秘密テキストの復号支援機能をperlで実装した。分散テキストの生成のための文データベースとしては、通信総合研究所の広報紙「CRLニュース」の20年分の全記事<sup>[12]</sup>を使用した。このデータベースを採用したのは、同じ分野の文によって構成されているという条件を考慮したためである。データベースのサイズは約5MBである。また、秘密テキストの復号支援については、形態素解析器「茶釜」<sup>[13]</sup>を用いた。

「盛岡駅前が会場です。」を秘密テキストとした場合に生成した分散テキスト(10枚)のうち、例として2枚を以下に示す。いずれも、自然言語テキストとして不自然ではないと思われる。

「最近、本来の電離層を介する伝搬よりも、むしろ宇宙通信に対して電離層が与える影響に関する研究の方が活発になっている傾向がある。もっとも内側の太線の円は衛星軌道を地球上に投影したものを表わすと同時に半径方向の臨界周波数目盛の0 MHzを示す。」

「この現象は雷放電による電波が電離層上部の多種類のイオンと作用し、特に重水素イオンと共鳴作用をすることによって生じたものと考え、これを重水素ホイッスラと呼ぶことにした。卒直に言って、当所は一般へのPRという面で、その施策が不十分であったと認めざるを得ない現況である。」

な	助動詞
数	名詞-一般
所	名詞-接尾-一般
古	接頭詞-名詞接続
年	名詞-一般
施	未知語
各	接頭詞-名詞接続
表	名詞-一般
も	助詞-係助詞
と	助詞-自立
つ	名詞-接尾-一般
目	名詞-固有名詞-地
山	
河	
域-	一般
程	助詞-副助詞
策	名詞-一般
字	名詞-接尾-一般
の	助詞-連体化
の	名詞-非自立-一般
題	名詞-一般
た	助動詞
盛	名詞-固有名詞-地
岡	
域-	一般
駅	
前	名詞-一般
が	助詞-格助詞-一般
会	名詞-一般
場	助動詞
で	記号-句点
す	助詞-連体化
。	名詞-一般
の	副詞-一般
己	接頭詞-名詞接続
よ	名詞-非自立-一般
に	助詞-格助詞-一般
不	助詞-自立
の	名詞-数
と	名詞-一般
あ	助動詞
る	未知語
0	名詞-数
雄	名詞-一般
り	助動詞
紹	未知語
十	名詞-数
雑	名詞-一般
し	助詞-自立
る	名詞-数
一	

また、分散テキストを重ね合わせて得られる文字列を形態素解析した結果の一部を上に表示。秘密テキストである「盛岡駅前が会場です。」の一節(括弧の部分)が、2文字形態素の連なりになっており、適切な閾値を設けることにより、高い精度で機械的に切り出せることが示唆されている。

### 3.4 考察

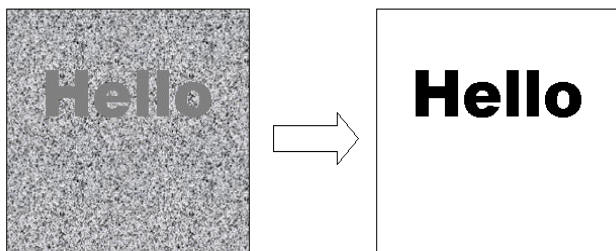
提案した手法では、形態素解析を援用して、秘密テキストをより視認しやすくする操作を施している。これは視覚

復号型秘密分散法において、分散画像の重ね合わせによりコントラストが低下した秘密原画像を、右下図のように、輪郭線処理などを援用してより視認しやすくすることに相当すると言える。

秘密テキスト自体が無意味な文字列（例えばパスワードなど）の場合には、形態素解析による復号支援処理は機能しない。その場合は、無意味な文字列の前後を意味ある文字列ではさむ等の対応をとる必要がある。視覚復号型秘密分散法の場合でも、無意味な画像を秘密原画像とした場合には、視認によって抽出することは困難である。

提案した手法は、分散テキストが欠けていても大まかに復号できてしまうという、秘密分散方式としてはかなり大きな欠点を抱えている。これは人間が自然言語テキストを読解する知的能力に関係することであるため、自然言語テキストを秘密分散法の情報隠蔽媒体とする場合につきまとう本質的な問題といえる。

重ね合わせの順序を入れ替えることで別の秘密テキストも隠蔽する方法については、今回は検討しなかったが、原理的には可能であり、今後の課題としたい。また3.1項の対処3で述べた、秘密テキストをちぎって散らして配置する方法についても検討したい。



## 4. ランダムデータの秘密分散

3節においては意味がある自然言語文をいくつかの意味のある自然言語文に分散して秘密に保持する方法を紹介した。次にランダムなバイナリデータを分散して自然言語の中に埋め込むことを考える。この場合のランダムデータは目的となるシステムへの秘密鍵ないしはパスワードとなるものである。よってそれ自体はセマンティクス（意味）を持たないバイナリ列である。よって、目的の秘密データに対して分散テキストを作成するのではなく、まず先に分散テキストを作成し、それから復号される自然言語文からランダムなバイナリデータを作成すれば良いだろう。実際にはテキスト秘密分散システムにおいて復号された自然言語文（のバイナリ表現）のハッシュ値を計算する。これはまったくランダムなものと考えられるので、このデジタルデータを目的のシステムへの秘密鍵として設定すればよい。ハッシュ値を取ることで、隠されている自然言語文の長さは任意でかまわない。それ以外は上記の方法で秘密の自然言語文を分散された自然言語文に隠蔽する。これにより、ランダムなデジタルデータでも自然言語の中に隠して秘密分散する目的が達

成できる。

## 5. ステガノグラフィとモデル化

テキスト秘密分散の既存の秘密分散に対する優位性は分散データが自然言語の中に隠されていることである。このことは分散データを自然な生活空間に埋め込むことを可能にして、分散データの取り扱い方を容易にする。計算機のハードディスクや電子的記録媒体ではなく、日記やノートにおいて日常生活空間に存在する普通の自然言語として分散された秘密が保持できる。これはステガノグラフィの効果を持っていると考えられる。以下にそれを説明しよう。

### 5.1 ステガノグラフィ

暗号システムにおいて暗号文を復号するためには秘密鍵が必要である。秘密鍵は通常自然言語として意味のないバイナリデータ（つまりランダムなバイナリ列）であるため、このことが逆に鍵の管理を困難にすることもある。自然言語文は通常的生活空間に氾濫している。一方、ランダムなバイナリデータはプログラムやデータとして計算機のハードディスクや電子的記録媒体に存在するデータがランダムに見えることがあっても、日常生活には現れてこない。ランダムなものが日常生活に必要な場面は何かのシステムのパスワードとしてであろう。パスワードのようなものでは個人の誕生日や電話番号と結びつけることにより記憶することも日常的に行なわれているかもしれないが、容易に推定することが出来るなど安全性に問題点があるためランダムなデータであるべきである。

通常秘密分散においては分散データはランダムなバイナリデータであり、もしくはOHPに印刷されたノイズ画像（visual secret sharing）である。これらのデータは計算機の中にファイルに保存されているか、もしくは物理的なデバイス（例えばOHPシート）に変形出来ない形で残されている。これらは、意味のある通常のテキストファイルや、意味のある画像や文字が印刷されたOHPシートと比べると、不自然なデータである。ここで不自然の意味が何であるかは明確に定義できるものではないが、ネットワークに何らかの手法で潜入した不正アクセス者がそれを秘密が隠されたパスワードのような重要な情報と認識する可能性は高い。分散された情報はそれ自体を入手してもそれだけでは隠された秘密情報を得ることが出来ないが、ネットワークセキュリティの観点からは不正者からの攻撃を避けるために分散情報そのものを隠してしまうことが望ましいだろう。

そこで、テキスト秘密分散のステガノグラフィ効果と有用性について考えてみたい。テキスト秘密分散のアイ

デアを誇張して図式的にあらわせば以下のように考えることができるであろう。

(1) テキスト秘密分散 =

秘密分散 + ステガノグラフィ + 自然言語処理

ここで等式(1)について説明しよう。テキスト秘密分散の復号:

$F(\text{分散情報1, 分散情報2, } \dots, \text{分散情報n})$   
= 秘密情報

においては、各分散情報(分散情報1、分散情報2、……、分散情報n)は意味のある自然言語文であり、秘密情報もまた意味のある自然言語文である。そして、秘密情報を抽出するアルゴリズム(または関数)Fは自然言語処理技術における形態素解析と自然言語文のもつセマンティクスを判断する何らかの閾値関数の総体として構成される。3節において我々は形態素解析により2文字以上の形態素の出現頻度による閾値の設定を行った。ある記号列が自然言語として認められるかどうか否かを判断するために我々は自然言語の持つ性質を利用している。自然言語においては2文字以上の形態素がおおくなる文字列は自然言語として意味をもつ確率が高くなる。この特徴を利用して閾値関数Fを構成することが出来る。分散データは日常生活の中に自然言語として埋め込まれていて、一見したところでは隠された情報を持っているとは見破られ難い。このことがテキスト秘密分散のステガノグラフィ効果である。等式(1)はテキスト秘密分散が自然言語処理技術を利用して、秘密分散にステガノグラフィ効果を加えた技術であることを示したものである。

ステガノグラフィ技術は秘密情報を他の意味のある情報または物理的なデバイスやメディアに秘密裏に埋め込む技術である。テキスト秘密分散において秘密情報は意味を持つ自然言語であることに注意されたい。

さて、それでは通常秘密分散では分散情報が自然言語文の中に埋めこめるかどうか考えてみよう。既存の秘密分散の復号法:

$F(\text{分散情報1, 分散情報2, } \dots, \text{分散情報n})$   
= 秘密情報

においてそれぞれの分散情報は無意味なもの(ランダムに選ばれたデータから加工されたもの)である。ここでFは分散情報から秘密情報を抽出するアルゴリズム、手段ないしは関数を表す。視覚復号型秘密分散法において各分散情報は意味のない画像データである。Shamirによる(k,n)閾値秘密分散法では秘密情報は意味のないバイナリデータであり、各分散情報はランダムに与えられた値の多項式関数の値である。このことから視覚復号型秘密分散法でもShamirの(k,n)閾値秘密分散法でも分散情報そのものに十分な意味を有した自然言語文を対応させることは困難である。もちろん、多項式関数の逆像をも

とめることが効率的に計算出来るならば可能であるが、一般的には困難である。

秘密分散における分散情報をどう管理するかという問題について、分散情報を暗号化して保存したほうが安全であるだろう。しかし、その場合においても暗号の秘密鍵と暗号化データは無意味なバイナリデータであり、それを保存された記憶媒体が盗まれた場合は、その暗号化されたデータは不正者になにか重要な情報を持っていると疑われる対象であることは間違いない。また、暗号化処理された鍵もまた無意味な(ランダムな)データであることから、無意味な(ランダムな)データを保存しておく行為がそのデータがなにか隠された重要な情報を有していることを指し示していると考えられる。バイナリデータを暗号化してなにか自然言語として意味のある暗号文が作成するようなことが出来れば、暗号文そのものがステガノグラフィとして暗号文であることを指し示す証拠を消し去ることが出来る。しかし、現実にはそのような自然言語的に無意味な(ランダムな)メッセージを意味のある暗号メッセージに変換する1対1写像なるものは存在しないだろう。また暗号文を復号するためには秘密鍵が必要である。しかし秘密鍵もまた自然言語的には意味のないバイナリデータであるため鍵の管理は簡単ではない。

一方、テキスト秘密分散において各分散情報は秘密情報とまったく異なる意味を持つ自然言語文であり、各参加者は意味のある分散情報を管理することが、意味のないバイナリデータの管理に比べて比較的簡単である。分散情報はそれ自体がセマンティクスを持つため、分散データ自体が漏洩した場合においても、それが隠されている秘密情報の分散情報であることが容易に判断されることはないであろう。秘密情報を抽出するために必要とされる情報がセマンティクスを持つので、それ自体を日常生活空間またはサイバー空間に埋め込むことが出来る。日常生活空間であれ、サイバー空間であれ、そのことを知らない部外者はセマンティクスを持つ何らかの自然言語であると理解して隠された秘密を抽出するための必要情報であると認識できないであろう。ネットワークセキュリティまたはコンピュータセキュリティの現状を鑑みれば、サイバー空間における情報が何らかの形で漏洩するかもしれないことは十分におこりえることである。現状の暗号技術がいかに安全であっても鍵そのものが漏洩するのであればネットワーク全体の安全性を保持できない。全体の安全性を高める努力として、暗号技術とは別に不正者の攻撃がたとえ成功してもステガノグラフィ技術により、結局求める分散情報がどれであるか推定できないことを目指すものである。この意味において、テキスト秘密分散は今までの秘密分散技術とまったくことな

る性質をもっているであろう。

## 5.2 モデル化

我々は簡略化した日本語文による実験を行った。しかし、テキスト秘密分散法の実現と安全性の解析のために必要である理論的なモデル化がまだ実現していない。自然言語の数理モデル化が理論的に行うことが出来れば、自然言語の情報について数理化が可能となり、情報通信セキュリティにおける応用が可能となる。

特に、現状では分散情報を作成することが極めて難しい。これは普通の秘密分散と比べた場合の大きな欠点となるかもしれない。分散情報の作成には自然言語処理技術、特に自然言語技術（または人工知能技術）の自然言語自動生成システム（談話生成）が必要になるかもしれない。自動的に自然言語文を創出することは人工知能における大きな課題である。

さらに、秘密分散法を考える上で大きな問題はアクセス構造をどのように与えるかである。現状の方法においては、秘密情報を得ることが許されるパーティはただ一通りに定まっているため、秘密分散の本質的な目的である柔軟なアクセス構造の構築について考察していない。テキストの文字列に順序があるために、アクセス構造として、順序構造を考慮した集合を考える必要があると考えられる。

自然言語の自動生成においては人工知能における談話生成技術が利用出来そうであるが、モデル化する場合にはむしろ Chomsky による生成文法の数理モデルを利用することも考えられるであろう。自然言語の深層部分の数理モデルとしての最も簡単なモデルである形式言語を数学的なモデルにして考えることは出来るであろう。

## 6. 謝辞

本研究のきっかけを与えて下さった、Pukyong National University の Prof. Ji-Hwan Park に感謝する。また、自然言語テキストを情報隠蔽媒体として適用する方法に関して、横浜国立大学の松本勉教授、東京大学の中川裕志教授、三菱総合研究所の村瀬一郎、井上信吾、牧野京子の各氏から有益な助言を賜っていることに感謝する。

## 【参考文献】

- [1] A. Shamir, "How to share a secret", Communications of the ACM, 612-613, 1979.
- [2] G. Blakley, "Safeguarding cryptographic keys", Proceedings of AFIPS National Computer Conference, 313-317, 1979.
- [3] M. Naor and A. Shamir, "Visual Cryptography", Advances in Cryptology-Eurocrypt'94, 1-12, 1994.
- [4] 加藤拓, 今井秀樹: "視覚復号型秘密分散法の拡張構成方式", 電子情報通信学会論文誌, Vol. J79-A, No. 8, 1344-1351, 1996年8月.
- [5] 有井幸太, 盛拓生, 坂井一雄, 今井秀樹: "積み重ね順

- 序を鍵とする視覚暗号方式", SCIS2000, 2000年1月.
- [6] 視覚復号型暗号製品「あわすとでーる」凸版印刷株式会社, <http://www.toppan.co.jp/aboutus/release/article463.html>, 2001年4月.
- [7] 松井甲子雄, "電子透かしの基礎", 森北出版, 1998年.
- [8] 中川裕志, 木村浩康, 三瓶光司, 松本勉, "辞書変換法に基づく日本語テキストへの情報ハイディング", 情報処論, Vol. 41, No. 8, 2272-2279, 2000年.
- [9] 松本勉, 中川裕志, 村瀬一郎, "ネットワーク向けインフォメーションハイディング技術開発 テキスト用フィンガープリンティング方式 FinPri.txt の開発", 情報処理振興事業協会 次世代デジタル応用基盤技術開発事業 先端の情報化推進基盤整備事業 論文集, 97-104, 2000年6月.
- [10] 滝澤修, "情報埋込・抽出方法及びその装置並びに記録媒体", 特願 2001-67597.
- [11] Moon-Soo Kim, Seong-Han Shin, Ji-Hwan Park, "New Construction for Multiple Visual Secret Sharing", SCIS2000, 2000年1月.
- [12] 通信総合研究所, "CRL ニュース", 創刊号~第238号, 1976年~1995年.
- [13] 奈良先端科学技術大学院大学情報科学研究科自然言語処理学講座(松本研究室), "日本語形態素解析システム茶釜 version 2.0 for Windows", 1999.
- [14] 滝澤修, 山村明弘, "自然言語文を用いた秘密分散の提案", コンピュータセキュリティシンポジウム 2001 343-348, 2001.