3-9 Secret Sharing Scheme Using Natural Language Text

TAKIZAWA Osamu, YAMAMURA Akihiro, and MAKINO Kyoko

Modifying the idea of the visual cryptography, we propose a method of sharing a secret key using natural language texts. Our target here is restricted to Japanese texts. Each participant obtains a share, which is a Japanese text in our scheme. When a certain number of participants retrieve the secret key, they supply their shares and pile up these natural language texts. The sequence of the first, second (and so on) letters occurred in the pile shows the secret text. The order of the pile is significant, and changing the order may yield the distinct secret text. It is easy to pile the shared natural language texts by computer operation. Human eyes can recognize the secret text from the piled texts, however, we aim to construct a natural language text secret sharing scheme employing a morphological analyzer because a meaningless phrase is a chain of morphemes consisting of one word with a high probability. We can make a shared natural text look like a natural text without any secret meaning by synthesizing using a text database.

Keywords
Information hiding, Text, Document, Secret sharing, Natural language processing

1 Introduction

Recent progress in computer and network technologies has led to an explosive increase in the distribution of digital content containing images, voice data, and text. As a result, the importance of information hiding, which embeds invisible information within content, is growing in a range of applications: assertion of copyright over digital content, identification of distribution routes, and as camouflage to prevent electronic eavesdropping over the information route.

Most traditional information hiding techniques throughout history have used natural language text as the cover media. A pixel in an image, for example, could correspond to a character in a secret text. Characters constitute a portion of the meaning of the text, and a character is linked in a fixed manner to a character code. Thus, if artificiality is added to the character codes to hide information in a text, this immediately affects the characters and their meanings. These effects may significantly damage the quality of the text and increase the risk that the artificiality is discovered. For this reason, to date most techniques for hiding information in text have included the data within the layout information—ultimately an image-based approach and thus inapplicable to plain text (with a limited number of exceptions)[1]-[4]. However, even in today's widespread multimedia environment, information exchange through text, as in email, remains the primary method of communication. The importance of text as a means of communication is not likely to diminish, and we can expect to see continued applications of information hiding using text as the cover media.

A secret sharing scheme is a technique that allows decoding of secret information only when shared pieces of information are combined[5][6]. As one implementation of such a scheme, Naor et al.[7] have proposed Visual

Cryptography (or the Visual Secret Sharing Scheme; referred to as VSSS hereafter). With this technique, the secret image appears only when two or more semi-transparent slides are superposed. This method has been extensively studied for research and commercialization as an information-hiding technique in which decoding does not require computers but instead relies on human observation[8][9][10].

This paper takes particular note of one characteristic of VSSS: it functions by superimposing image content—actual media[11]. To provide similar means for content other than images, we propose our Text Secret Sharing Scheme (referred to as TSSS hereafter), which uses natural language texts as the cover media. The proposed TSSS technique overlays two or more shared texts, and the sequence of characters read from the top layer to the bottom layer forms the secret text. From the sequence of characters obtained in the compiled layers, the secret text is extracted through morphological analysis.

Section **2** discusses the principles of TSSS relative to VSSS. Section **3** explains the method of generating the shared texts and implementing the technique. Section **4** discusses the set of assumptions required to extract and validate the assumptions. Section **5** discusses the requirements for generating shared texts that appear natural, and presents possible improvements to the proposed TSSS technique. Section **6** consists of a discussion. Section **7** describes future perspectives.

2 Principles of TSSS

TSSS can be defined as an information-hiding method in which secret text is divided into two or more shared texts; the secret text is recovered by superposing these texts. With the "skytale cipher", used by the ancient Greeks, the sender of the message wrapped a paper tape around a drum of a certain thickness, wrote on the paper in the direction of the drum's axis, and then unwound the paper tape to send to the recipient. The recipient wrapped the paper tape around another drum of the same diameter to recover the message. In other words, the skytale cipher scrambles the character sequence at a constant interval according to the diameter of the drum. TSSS applies this concept of the skytale cipher as a process corresponding to the superposition in VSSS. Specifically, two or more shared texts are each written horizontally; the first character of the texts are aligned with each other (under the assumption that the character widths are the same), and the secret text appears in the vertical sequence of the characters at a certain position. Figure 1 shows an example. In this case, each turn of the paper tape in the skytale cipher corresponds to each shared text (each line). In this example, it should be noted that each shared text has a self-contained meaning.

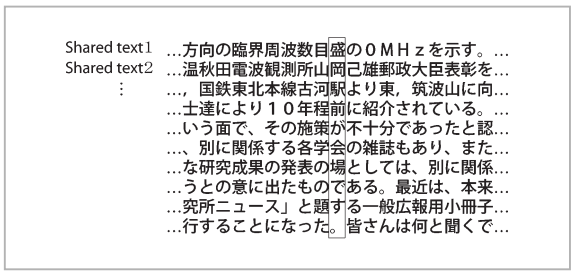

Fig. 1 *Principles of TSSS*

The character sequences obtained in the layered shared texts are meaningless except for the embedded secret text, which holds the intended meaning. In the Japanese texts shown Fig.1, the character sequences obtained in the pile are read vertically, for example, as 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皆, and ０雄り紹十雑しる一さ. The secret text in this case is 盛岡駅前が会場です。(The meeting place is in front of Morioka Station.) To find the secret text, natural-language recognition is required. In other words, the secret text is distinguished from other text in that the former conveys natural-language meaning.

The principles of TSSS described above may be stated as follows: the number of characters from the beginning of the shared text up to the first character of embedded secret text is

required to be the same for all shared texts; this requirement is used as a clue in extracting the secret text. Generally, if the positions of all the characters constituting the secret text in each shared text is known at extraction, the number of characters in each shared text does not need to be the same. However, when the number is not the same, it is necessary to stipulate that positional information—specifically, the key—is shared both by the embedding agent and the extractor of the secret text. The TSSS proposed in this paper sets a high priority on reducing the frequency of key use, as in VSSS. Thus, each shared text in TSSS features the same number of characters prior to the initial character of the secret text.

In the technique proposed by Shamir and Blakley et al., the shared data are simple random bit sequences, and the data is stored on media such as hard disks. As the shared data is itself random, there is the risk that a third party with access to the storage medium will clearly see that the text contains secret information. Particularly in text data, few documents consist of random character sequences. Such a document immediately raises the suspicion of hidden information, and this suspicion represents a threat to the effectiveness of information hiding. Thus, the shared text should form a natural document with meaning, to avoid eliciting attacker suspicion concerning the text in question.

Based on the above principles, Section **3** explains the method of generating the shared texts and the implementation of this technique, and Section **4** discusses the set of assumptions required to extract and validate the assumptions.

3 Generation of shared texts

As discussed in the previous section, our aim is to prepare shared texts that consist of meaningful natural language. However, an enormous amount of knowledge and complicated algorithms are required for a computer to synthesize such meaningful natural language through individual word combinations. As the content of the shared text does not need to convey specific information in TSSS, but rather must simply appear natural, it is inefficient to synthesize texts from individual words. We thus propose that text be stored in a database not by word but by sentence (ending with a period), and that shared texts be generated through the connection of these sentences.

The following shows the algorithm for generating the shared texts.

[Definition]

The symbol { } denotes a set, and the symbol < > denotes an ordered set. An uppercase alphabetic character denotes a text (or a set of texts) and a lowercase alphabetic character (except suffixes) denotes a character.

Secret text E is defined as an ordered set consisting of the character sequence $<e_1, e_2, ..., e_\varepsilon>$. Here, ei is a Japanese character and E is a Japanese sentence containing ε characters. For the example of Fig.1,

E = < 盛, 岡, 駅, 前, が, 会, 場, で, す, 。>.

In this example, ε is 10.

The text database D is expressed as a set of texts, $\{T_1, T_2, ...\}$.

An element of D, T_x, is an ordered set consisting of the character sequence $<t_{x_1}, t_{x_2}, ...>$. The last element of T_x (the last character) is the Japanese period, "。". There are no restrictions regarding the length of T_x.

In the following, we consider a case in which ε is the same as the number of shared texts.

[Process procedure]

(1) Extract texts, each containing a character in the secret text, from the database.

For all i that satisfy $1 \le i \le \varepsilon$, extract text T_{x_i} containing e_i from D. Let $e_i = t_{x_i z_i} (\in T_{x_i})$. As a result,

$$<e_1, e_2, \cdots, e_\varepsilon> \equiv <t_{x_1 z_1}, t_{x_2 z_2}, ..., t_{x_\varepsilon z_\varepsilon}>$$

(2) Process for matching the number of characters

Next, extract $\{T_{w_i}\}$ $(\in D, 1 \le i \le \varepsilon)$ that satisfy

$$T_{w_i} = <t_{w_i 1}, t_{w_i 2}, ..., t_{w_i y_i}>$$

and

$$y_1 + z_1 = y_2 + z_2 = ... = y_\varepsilon + z_\varepsilon$$

from D. Let this value be denoted as j. Here, each T_{w_i} consists of one or more sentences.

(3) Synthesis of shared text

Let the shared texts $<S_1, S_2, ..., S_\varepsilon>$ be denoted as

$S_1 = <T_{w_1}, T_{x_1}>$
$= <t_{w_1 1}, t_{w_1 2}, ..., t_{w_1 y_1}, t_{x_1 1}, t_{x_1 2}, ...>,$
$S_2 = <T_{w_2}, T_{x_2}>$
$= <t_{w_2 1}, t_{w_2 2}, ..., t_{w_2 y_2}, t_{x_2 1}, t_{x_2 2}, ...>,$
...
$S_\varepsilon = <T_{w_\varepsilon}, T_{x_\varepsilon}>$
$= <t_{w_\varepsilon 1}, t_{w_\varepsilon 2}, ..., t_{w_\varepsilon y_\varepsilon}, t_{x_\varepsilon 1}, t_{x_\varepsilon 2}, ...>.$

(End of process)

The principle of the above process is to synthesize $\{S_i\}$ for all i that satisfy $1 \le i \le \varepsilon$, so that the j-th character of S_i is e_i (= $t_{x_i z_i}$). $\{T_{w_i}\}$ is inserted only for matching the number of characters.

The above procedure is implemented via perl script. All articles[12] in the past 20 years of "CRL News", the newsletter of the CRL (Communications Research Laboratory), are used as the text database. All of these articles contain technical information limited to nearly a single field—communication technology—and thus lend themselves to combination. The size of the database is approximately 5 MB.

Among the 10 shared texts generated with the secret text, 盛岡駅前が会場です。(ε = 10), two examples are shown below. Of the characters that constitute the secret text (referred to as the secret characters), the first example contains 盛 and the second example contains が.

最近は、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０ＭＨｚを示す。

(The recent tendency shows that more studies are conducted in relation with the effect of the ionosphere on space communication than the communication by the ionosphere itself (forward propagation ionospheric scatter: FPIS). The innermost thick circle represents a projection of the satellite orbit onto the earth and indicates the 0 MHz position of the critical frequency scale in the radial direction at the same time.)

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。(Considering that this phenomenon is caused by the interaction between radio waves due to lightning discharge and many types of ions in the upper ionosphere, particularly through the resonance with deuterium ions, we have decided to call the phenomenon the deuterium whistler. Frankly, the present status of CRL shows that we have not conducted sufficient activities in the light of general public relations.)

As an example of the operation of the algorithm, the process for generating the shared text above containing 盛 is described below.

(Step 1)

Find a sentence containing the secret character 盛 from the text database and extract the following sentence, T_{x_i} (T_{x_i} is referred to as the extracted sentence.)

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０MHzを示す。 (The innermost thick circle represents a projection of the satellite orbit onto the earth and indicates the 0 MHz position of the critical frequency scale in the radial direction at the same time.)

Here, if two or more sentences containing the same secret character are present in the database, the sentence registered earlier in the database is selected as the extracted sentence. If a sentence containing the secret character is not present in the database, the process fails and ends. If the process fails, modify the secret text manually or take other

corrective measures, and repeat the process.
(Step 2)

Find extracted sentences similarly for the remaining secret characters.
(Step 3)

When extracted sentences are obtained for all secret characters, find the sentence containing the largest number of characters between the beginning of the sentence and the secret character and count the total number of characters in this sentence (referred to as the maximum number of characters). In the above example, the extracted sentence for 岡 provides the maximum number of characters, 110.
(Step 4)

For the rest of the extracted sentences, subtract (i) the number of characters from the beginning of the sentence to the secret character from (ii) the maximum number of characters. For the extracted sentence containing 盛, there are 47 characters from the beginning of the sentence to the secret character; thus, the calculated difference here is 63.
(Step 5)

Extract a sentence, T_{w_i}, of a length equal to the calculated difference from the text database. If two or more sentences of the same length are present in the database, select the sentence registered earlier in the database. For the extracted sentence containing 盛, the following sentence is extracted as the 63-character T_{w_i}:

最近は、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。(Recent tendencies show that more studies are conducted in relation with the effect of the ionosphere on space communication than the communication by the ionosphere itself (forward propagation ionospheric scatter: FPIS).

(Step 6)

Generate the shared text S_i by connecting T_{w_i} and T_{x_i}.

This is the end of the process.

4 Extraction of secret text

To extract the secret text by compiling the shared texts, the embedded position of the secret text must be identified. If the information regarding the embedded position is not available at extraction, the secret text and other character sequences must be separated based on natural-language properties. For this purpose, the secret text must be a phrase with meaning. This textual restriction recalls the difficulty in VSSS of extracting the original secret image from background visually when the secret image is meaningless, as the outline of the image cannot be identified. Thus, the corresponding textual restriction is reasonable in the application of a secret sharing scheme. With this restriction, the possibility that character sequences other than the secret text will accidentally hold meaning as phrases is small. Thus, extracting a meaningful phrase through natural language processing is considered equivalent to extracting the secret text. Accordingly, the secret text can be extracted by visual observation. However, here we investigate extraction using natural language processing.

Based on the above premises, let us establish the following assumption in order to extract the secret text through natural language processing.
[Assumption]

Morphological analysis, which divides a text into morphemes (the minimum unit constituting a word), is a basic type of natural language processing. In morphological analysis, meaningless phrases frequently form a chain of single-character morphemes.

To verify the above assumption, morphological analysis is performed for sentences with meaning and random character sequences, and the appearance frequencies of single-character morphemes are compared.

8098 Japanese character sequences with meaning are extracted from main technical expository texts[13] and random character sequences are generated from the same texts*. These character sequences are then morpho-

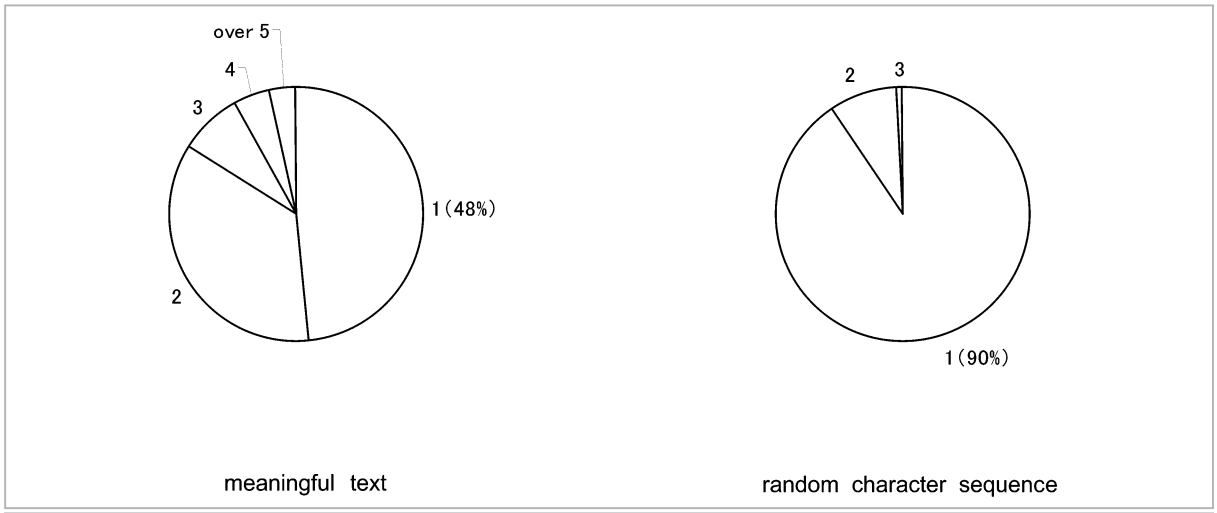

Fig.2 *Relationship between the number of characters in a morpheme and the frequency of appearance*
The numerical figure is the number of characters in each morpheme.

logically analyzed. Here the "Chasen"[15] morphological analyzer and its standard morphological dictionary are used. The number of morphemes in character sequences with meaning is 4,444, while the number of morphemes in the random character sequences is 7062. Figure 2 shows the relationship between the number of characters in a morpheme and the appearance frequency of each character. In the technical expository text, single-character morphemes comprise slightly less than half of total morphemes (48%), while they constitute 90% of the total for random character sequences, with the longest morpheme containing only three characters. Next, Figure 3 shows the relationship between the length of a single-morpheme chain and the frequency of appearance. In technical expository text, the single-character morphemes constitute approximately half of the total, as shown in Fig.2, while single-character morpheme chains containing three or more single-character morphemes form only 11% of the total. On the other hand, for random character sequences, these long chains constitute 86% of the total. From these results, we can conclude that a phrase with meaning can be extracted at high probability by locating short chains of single-character morphemes, which confirms the validity of the established assumption. In implementation, phrases for which Chasen produces a single-character morpheme chain containing less than three single-character morphemes are considered as candidates for the secret text. With this threshold value, the error rate in extracting a random character sequence as the secret text, in other words, "1—[compatibility]" can be estimated as 14%, and the error rate in overlooking a phrase with meaning ("1—[reproducibility]") can be estimated as 11%. If the threshold length of the single-character morpheme chain is larger, reproducibility increases but compatibility decreases. If it is shorter, the reverse occurs.

Figure 4 shows the results of morphological analysis by compiling the shared texts shown in Fig.1 in the correct order. The phrase with less than three single-character morphemes in a chain 盛岡駅前が会場です。 (the part indicated in the parentheses), is extracted as candidate secret text.

* BookNoise ver1.01[14], free software for encoding and decoding character strings, was used. Reference[13] is encoded with this tool once (into an ASCII character sequence) and converted to a completely different Japanese character sequence with a different key. We consider the resultant character sequence as randomized, although not a random character sequence in a strict mathematical sense.

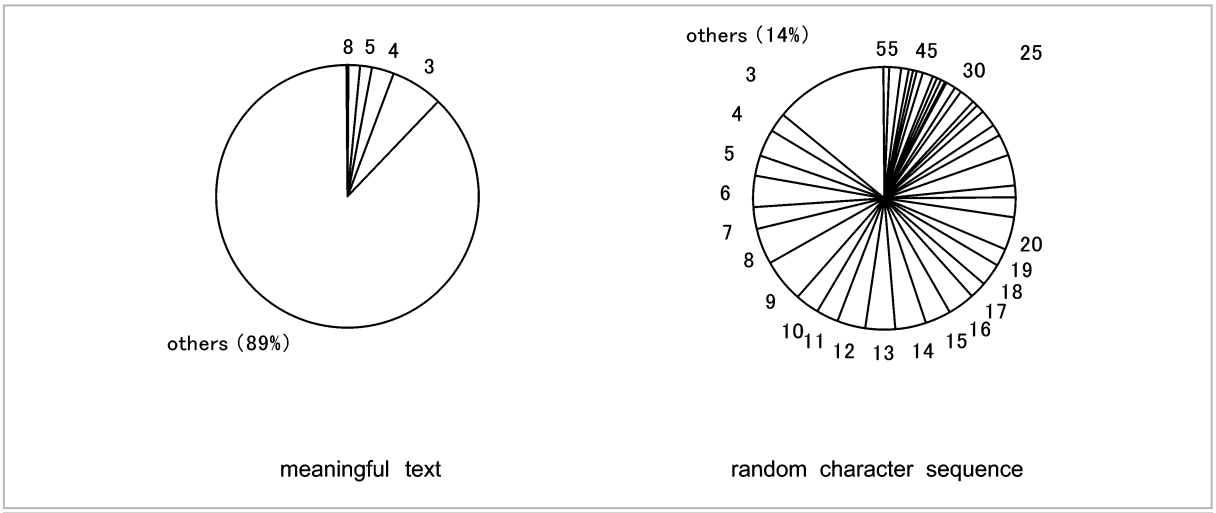

Fig.3 *Relationship between the length of a single-character morpheme chain and the frequency of appearance*
The numerical figure is the number of single-character morphemes in each chain.

5 Discussion of the naturalness of shared texts

In a secret sharing scheme, the secret information cannot be recovered completely if some of the shared text is lost. Thus, the threat is that a third party may discover that information is concealed within the shared text and may destroy or modify the data. To make the secret information difficult to detect, it is important to maintain naturalness in the generated shared texts.

The proposed TSSS technique does not contain a process for synthesizing the sentences themselves. The shared texts are generated by connecting two or more sentences extracted from the text database consisting of existing human-made sentences. Thus, the proposed technique does not present a problem of naturalness in terms of each standalone sentence. The problem lies in maintaining naturalness in the connections between the sentences.

The naturalness in the connections between sentences depends on whether the sentences are consistent with respect to a single topic; in other words, whether they have a unified meaning as a whole. Yamamoto et al. showed that "cohesion" between sentences can be used for quantitative assessment of the meaning of a text[16]. According to these authors, cohesion is expressed in terms of the connecting words, such as conjunctions and adverbs, and words with similar meanings (repeated words, hypernyms or hyponyms, synonyms, and antonyms). In other words, if two sentences are connected by a conjunction or if words with similar meanings appear frequently, the sentences have strong cohesion. Thus, the naturalness of the shared texts can be improved by using these proposed criteria of coherence to select the extracted sentence when two or more candidate sentences are present in the database. While the current method selects whichever sentence was first registered, the improved method employs the following process.

For the sentence T_{x_i} containing the character constituting the secret text (the secret character), and the sentence T_{w_i} to identify the number of characters from the beginning of the text to the secret character, calculate the cohesion based on the method proposed by Yamamoto et al. throughout $1 \leq i \leq \varepsilon$. Select the combinations of T_{x_i} and T_{w_i} that maximize the sum of cohesions through $1 \leq i \leq \varepsilon$ as the shared text.

In the proposed technique, the content of each shared text is arbitrary as long as textual naturalness is maintained. Thus, the text data-

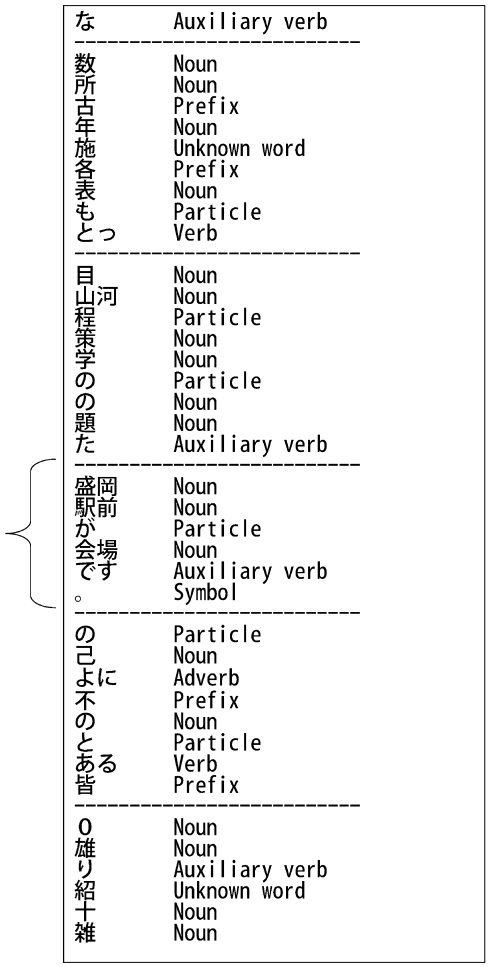

Fig.4 *Results of morphological analysis for the shared texts in Fig.1*

base may be selected arbitrarily without restrictions on content. The improved cohesion through the above process increases the naturalness of the shared text; moreover, replacing the text database can further improve the naturalness of the shared text if the same method is employed. For example, with a text database consisting of sentence data divided into detailed genres or categories, restrictions may be set stipulating that sentences belonging to the same sub-category should only be used for T_{x_i} and T_{w_i}; this will increase the naturalness of the generated shared text even further.

As such, the proposed technique allows for expansion to improve the naturalness of the shared texts, as is appropriate for an initially proposed text secret sharing scheme.

6 Discussion

Except for methods that use the order of superposition as a key, many VSSS methods use simple superposition of shared data upon extraction without a key. On the other hand, the currently proposed technique uses the order of compiled shared texts as the key and assumes that this key is shared at embedding and extraction. If the compiling order is incorrect, visual extraction will not be able to find the phrases with meaning. The technique proposed in this paper, on the other hand can find the secret text in principle, even in the absence of information on superposition, using all pos-

sible layering permutations in its analysis. The number of permutations for the compiling order is $\varepsilon!$; thus, the computational cost increases if ε is large, which presents a problem. Nevertheless, by speeding up the morphological analysis and applying parallel processing, it is possible to implement an extraction method without using a key, as in VSSS.

The proposed technique requires that the number of characters in the secret text be the same as or smaller than the number of characters in the shared texts. This means that many shared texts are required when embedding a long secret text. Thus, the proposed technique is not suitable for applications requiring exceedingly long secret texts. Use of the technique is limited by the requirements of the secret information. Nevertheless, this limitation is equivalent to the difficulty in VSSS of using complicated or low-contrast images as the secret image. Such restrictions are inevitable when applying a secret sharing scheme to content. Such a scheme is useful when distributing secret information for key recovery and key escrow and is essential when implementing a multi-party protocol. The technique can also be used for multiplexing a cipher to increase security in cipher communication using two or more cipher techniques. On the other hand, the proposed technique is not suitable for sharing long secret texts, as already discussed; it is most suitable for applications involving distributed sharing of keys.

An ordinary secret sharing scheme cannot restore the secret data if even a single item of shared data is missing. On the other hand, the proposed Text Secret Sharing Scheme increases in completeness as the number of missing shared texts decreases. Thus, when enough shared texts are collected to interpolate the missing characters of the secret text based on context, the secret text is essentially recovered even if not all shared texts are present. This is inevitably the case as long as the shared data and the secret data both consist of natural-language texts with meaning. However, the proposed technique uses morphological analysis to extract the secret text, so that splitting of morphemes is highly possible when a portion of the secret characters is missing. This property is expected to lead to failure in extracting the missing secret text portion as a phrase with meaning, and decoding becomes highly unlikely. The relationship between the deficiency of the secret text and the security in the recovery of the secret text will require future quantitative verification.

7 Conclusions

This paper proposes a secret sharing scheme that uses a natural-language text as the medium for information hiding. The paper also discusses the results of implementation of a shared text generation function and a secret text extraction function. For the latter, morphological analyses are performed for character sequences with meaning and for random character sequences, and the results are compared. As a result it is demonstrated that a phrase that produces three or less single-character morphemes in a chain represents an appropriate threshold. The naturalness of the generated shared texts is also discussed, with the conclusion that the proposed technique may be improved using the concept of cohesion between sentences constituting each shared text.

In the future, studies will mainly focus on resolving the problems pointed out in the discussion above and on evaluating the naturalness of the generated shared texts, including subjective evaluation experiments. We will also consider properties of a text database suitable for the proposed technique and evaluate these properties through application to a large number of sample secret texts. Additionally, we will study specific applications of the proposed technique.

Acknowledgments

We would like to express our gratitude to the late Prof. Ji-Hwan Park of PuKyong National University, who served as an inspiration in this study. We also thank Prof. Tsutomu

Matsumoto and the members of his laboratory at Yokohama National University; Prof. Hiroshi Nakagawa of the University of Tokyo; and the members of the Mitsubishi Research Institute, Inc. for their useful advice.

References

1 M.J.Atallah, V.Raskin, C.F.Hempelmann, M.Karahan, R.Sion, U.Topkara, and K.E.Trizenberg, "Natural Language Watermarking and Tamperproofing", Proc. Int. Workshop IH 2002, LNCS 2578, pp.196-212, Springer, 2002.

2 Hiroshi Nakagawa, Hiroyasu Kimura, Koji Sanpei, and Tsutomu Matsumoto,"Information Hiding for Japanese Text Based on Replacing Words with Dictionary",IPSJ Journal, Vol.41, No.8, pp.2272-2279, 2000. (in Japanese)

3 Tsutomu Matsumoto, Hiroshi Nakagawa, and Ichiro Murase, "Information hiding technical devel-opment for network-development of finger printing system for document -FinPri.txt", Information-Technology Promotion Agency, 2000.(in Japanese).

4 Osamu Takizawa, "A method of embedding and extracting information, its equipment, and recording medium", Unexamined Patent Publications 2002-269074. (in Japanese)

5 A.Shamir, "How to share a secret", Communications of the ACM, pp.612-613, 1979.

6 G.Blakley, "Safeguarding cryptographic keys", Proceedings of AFIPS National Computer Conference, pp.313—317, 1979.

7 M.Naor and A.Shamir, "Visual Cryptography", Advances in Cryptology-Eurocrypt'94 , pp.1-12, 1994.

8 Taku Katoh and Hidei Imai, "An Extended Construction Method of Visual Secret Sharing Scheme", IEICE Trans., Vol.J79-A, No.8, pp.1344-1351, 1996. (In Japanese)

9 Kota Arii, Takuo Mori, Kazuo Sakai and Hideki Imai, "Stacking-order-key Visual Cryptography", The 2000 Symposium on Cryptography and Information Security, B46, IEICE, 2000. (in Japanese)

10 "Awasu-to-deeru -A material of Visual Secret Sharing Scheme", Toppan Printing Co. Ltd., http://www.toppan.co.jp /aboutus/release/article463.html, 2001. (in Japanese)

11 Osamu Takizawa and Akihiro Yamamura, "Secret Sharing Scheme Using Natural Language Text",IPSJ Journal, Vol.45, No.1, pp.320-323, 2004. (in Japanese)

12 "CRL News", No.1-No.238, Communications Research Laboratory, 1976-1995. (in Japanese)

13 Hirosho Nakagawa, Osamu Takizawa, and Shingo Inoue, "Information Hiding on Digital Documents", IPSJ Magazine, Vol.44, No.3, pp.248-253, 2003. (in Japanese)

14 "BookNoise ver.1.01", http://www.vector.co.jp/soft/win95/util/se267011.html

15 "ChaSen -A morphological analysis system", version 2.0 for Windows, Computational Linguistics Laboratory, Graduate School of Information Science, Nara Institute of Science and Technology, 1999. (in Japanese)

16 Kazuhide Yamamoto, Shigeru Masuyama, and Shozo Naito, "Cohesion Structure of Japanese Sentences and Paragraphing", IPSJ Journal, Vol.35, No.10, pp.2029-2037, 1994. (in Japanese)

***TAKIZAWA Osamu**, Ph.D.*

Senior Researcher, Security Advancement Group, Information and Network Systems Department

Contents Security, Telecommunication Technology for Disaster Relief

MAKINO Kyoko

Mitsubishi Research Institute, Inc.

Information Security

***YAMAMURA Akihiro**, Ph.D.*

Group Leader, Security Fundamentals Group, Information and Networks Systems Department

Information security, Cryptography, Algebraic systems and their algorithms