Reversible Text Watermarking Algorithm Using Prediction-error Expansion Method

Wenbin Fei[1], **Xianghong Tang**[1,2]

[1]School of Communication Engineering Hangzhou Dianzi University, Zhejiang Province, China
[2]School of Information Engineering Hangzhou Dianzi University, Zhejiang Province, China
feiwenbin111@sina.com, tangxh@hdu.edu.cn

Abstract - In order to avoid the permanent change of the text content caused by watermark embedding, this paper proposes a reversible watermarking algorithm for Chinese text document using prediction-error expansion. The algorithm takes the sentence as the unit, through the context collocation degree to select the words that can be replaced, using prediction-error expansion and chaos sequence to realize embedding. Results show that this algorithm not only has the higher security, but also can extract watermark effectively, and restores the original text intact.

Index Terms - reversible watermarking; prediction-error expansion; text document; context collocation degree; chaos sequence

1. Introduction

Digital watermarking technology uses random redundancy of carrier data, embedding the secret information in the data carrier, making it not be found by people's perception system, and between before and after of the watermark algorithm, the distortion usually can't be perceived by visual sense. This method is putting people in a distorted constraints of the human eyes which can't detect in small scope. But in certain application areas, the carrier data is very sensitive to changes, don't allow carrier data permanent loss, such as medical, military field and judicial field. The change of the carrier data caused by embedding watermark makes the quality of carrier data seriously reduce, so a kind of technology called a reversible watermark technology in recent years received extensive attention. Reversible watermark requires decoder can not only watermark extraction, but also restore the original text intact [1].

At present, the researches of reversible watermark are mainly concentrated in the image field [2-6], but for text document which is one of the most widely applied has little research reports. Tian first proposed the reversible algorithm based on pixel difference expansions(DE) in 2003 [2]. The algorithm used the pixels value between two adjacent pixels to extend, embedded the watermark to the least significant bit(LSB), realized the blind extraction and pixel recovery. After that, the pixel difference expansions are widely used in image reversible algorithm. Thodi proposed another reversible algorithm based on prediction-error and expanding(PEE) in 2004 [3]. This algorithm was improved in the embedded capacity and the image fidelity. In the reversible text watermarking algorithm researches, Liu Zhi-jie proposed two reversible text watermark algorithms using integral reversible transform and an improved difference expansion method based on image scheme [7]. Two schemes were adopting related principles of integer reversible transformation and difference expand in the reversible image watermarking algorithms, through the synonym substitution way to embed watermark, but in those algorithms hadn't considered the distortion of the text semantic after algorithm, text semantic would produce distortion and deviation. Jiang Chuan-xian proposed a robust reversible text watermark algorithm [8], this algorithm had strong robustness, through calculating the frequency of the occurrence probability of common words collocation appears in the training corpus, using the original word XOR the max occurrence probability creating recovery data in embedding algorithm, but the extraction and reversible algorithm used a large number of redundant information(recovery data) to restore the original data.

Therefore, in view of the shortcomings of the above reversible text watermarking algorithm, this paper firstly calculates the context collocation degree(CCD) by the context of the sentence words, to get the synonyms of the original text word, which doesn't produce semantic deviation, and then, uses the satisfying words to expand prediction-error. By using chaos mapping, realizes 1-bit watermarking information embedded by every time of synonyms replacement and also improves the security of the watermarking algorithm. Through the spillover information generated by chaos mapping can completely restore the original data in the recovery algorithm, discussing a reversible text watermarking algorithm using prediction-error expansion.

2. Context Collocation Degree and Prediction-error Expansion

A. Context Collocation Degree

Since the principle of proposed watermarking embedding algorithm is by using the replacement of synonyms method to realize watermarking embedding. Firstly, we need to find the words which have synonyms in the text, and then use the words' synonyms to replace themselves. The algorithm must consider the language environment of the context to avoid the replacement causing semantic chaos [9, 10], this means that not all synonyms can be replaced. To solve this problem, the paper uses the context collocation degree algorithm [11] to eliminate the context semantic ambiguity by synonyms replacement.

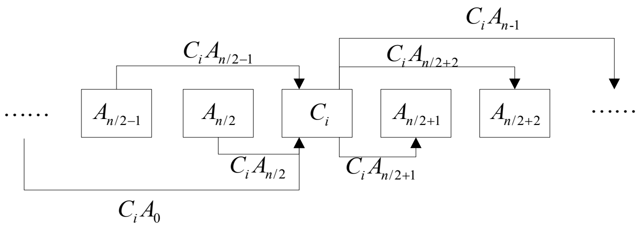


Fig. 1 The synonym and its context words in the text

In order to confirm the relationship of the collocation among words in context, this paper uses the Chinese Language Technology Platform(LTP) made by Harbin Industrial University to grammar analysis [12], to get the relationship of the collocation between word C_i and other words $A_j(j=0,1,...,n-1)$ in every sentence, shown in Fig. 1, and through (1) to compute the words' CCD.

$$CCD(C_i;A_0,A_1,...,A_{n-1})=\sum_{j=0}^{n-1}CD(C_iA_j) \quad (1)$$

In (1), $CD(C_iA_j)$ is the frequency of collocation degree(CD) of two words, C_i and A_j. Its value taken from the SougouR made by Sougou Labs [13]. Therefore, we know the CD of all C_iA_j, through (1), we can calculate the CCD of word C_i. In original sentence, using the synonymous $C_i^z(z=0,1,...,m-1)$ of C_i replacing C_i, we also can get word C_i^z's CCD.

B. Prediction-error Expansion and Chaos Mapping

If the initial value of a variable is x, the predictive value is $\hat{x}$, through (2), we can realize 1-bit watermark information w be embedded [3].

$$\begin{cases} p_e=x-\hat{x} \\ p_e'=2p_e+w \\ x'=\hat{x}+p_e'=2x-\hat{x}+w \end{cases} \quad (2)$$

Where p_e is the prediction-error, p_e' is the prediction-error after watermark information embedded, and x' is the predicted value.

In the prediction-error expansion algorithm, the value of x' may be beyond the scope of variables x, called overflow and underflow conditions. Since text document has less embed redundant, it can’t store overflow and underflow information directly. Therefore, in order to overcome problems, we need convert them into binary sequences by using the Logistic chaos sequences [14].

$$y_{i+1}=4y_i(1-y_i)(0\le y_i\le 1) \quad (3)$$

Through (4), generates a pseudo random binary sequence.

$$s_i=\begin{cases}1, if\, 0.5<y_i\le 1 \\ 0, if\, 0\le y_i\le 0.5\end{cases} \quad (4)$$

In the Logistic mapping processing, the pseudo random binary sequence is closely related to the initial value of chaos mapping. In order to make the Logistic binary sequence match the sequence of overflow and underflow sequence, we adopt the chaotic search method, the steps as follows:

Firstly, set a small initial value y_0, to produce a chaotic binary sequence with length N, such as $y_0=0.0001$, N=10000. If the generated sequence doesn't contain the overflow and underflow sequence, then the initial value will add a small step, such as $y_0=y_0$+0.0001, get chaos sequence again until the chaotic binary sequence match the overflow and underflow sequence. If the overflow and underflow sequence is too long, it can be cut to section sequence, then map sub-sequence with the same mapping method.

Through this method, overflow and underflow information will match to chaotic sequence, recording the initial value y_0 and starting position B while overflow and underflow sequence matching to chaotic sequence. At receiver, we can use the secret key y_0 and B to get the overflow and underflow information in the watermark recovery algorithm.

3. Algorithm Principle and Realization

Based on the above analysis, the basic principle of the reversible text watermarking algorithm in this paper is, through calculating the CCD of the sentence words to get the non-deviation of synonyms in semantics, using each word that satisfy the threshold conditions to calculate prediction-error expansion. Making use of mapping to realize watermarking embedding via synonyms replacement, and improve the security of embedding watermark.

A. Embedding Algorithm

The processing of embedding algorithm is shown in Fig. 2. The implementation steps as follows:

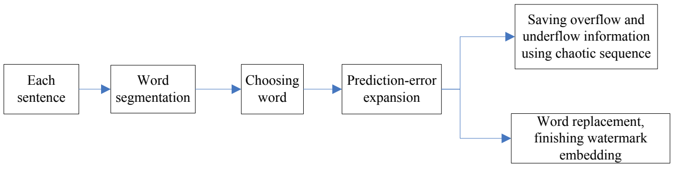

Each sentence
Word segmentation
Choosing word
Prediction-error expansion
Saving overflow and underflow information using chaotic sequence
Word replacement, finishing watermark embedding

Fig. 2 The processing of embedding algorithm

1) Word segmentation. Getting the collocation relationship among words in context from LTP.

2) Choose words. Through (1), calculating the CCD, locating the candidate replaceable words and their candidate replaceable synonyms, preparing for the prediction-error expansion method. Specific procedure is as follows:

a) According to the collocation relationship among words in context, calculating the word C_i 's CCD. In order to improve the security of the algorithm, set a threshold T , finding the word C_i that its CCD is greater than the threshold T . The word C_i is the candidate replaceable word.

b) Finding the synonyms $C_i^z(z = 0,1,...,m-1)$ of the each candidate replaceable word C_i , in the synonym dictionary. Using each synonym C_i^z to replace the original words in the text, calculating the C_i^z 's CCD, finding the candidate replaceable synonyms $C_i^{z'}$ $(z' = 0,1, 2, \cdots, M-1)$ that their CCD are greater than the threshold T . Finishing the screening of the synonyms.

c) According to the value of CCD in words $C_i^{z'}$ and C_i , sorting them from small to large, generating the new group of each locating word $NC_i^{z''}$ $(z'' = 0,1, 2, \cdots, M)$.

3) Embedding watermark using the prediction-error expansion method and classifying the overflow and underflow information.

Suppose order number z" of each candidate replaceable words C_i in group $NC_i^{z''}$ is the initial value x . M is predicted value $\hat{x}$, and the watermark information is w . According to (2), calculating x' , $x' \in [-\hat{x}, \hat{x}+1]$. And then convert the overflow and underflow information into binary sequence according to ***Rule 1***.

Rule 1:

a) when $x' \in [0, \hat{x}]$, x' will not overflow or underflow, denoted as Class I, with the representation of binary '00';

b) when $x' \in [-\hat{x}, 0)$, x' will generate underflow, denoted for Class II, with the representation of binary '10';

c) when $x' \in (\hat{x}, \hat{x}+1]$, x' will generate overflow, denoted for Class III, with the representation of binary '11'.

4) According to ***Rule 2***, calculating x' for x''($x'' \in [0, \hat{x}]$).

Rule 2:

a) if x' belong to Class I, which means $x' \in [0, \hat{x}]$, does not make any change, $x'' = x'$;

b) if x' belong to Class II, which means $x' \in [-\hat{x}, 0)$, calculating its absolute value, $x'' = -x'$;

c) if x' belong to Class III, which means $x' = \hat{x}+1$, calculating $x'' = 2\hat{x} - x' = \hat{x} - 1$.

5) Finding out the corresponding word $C_{x''}$ in $NC_i^{z''}$ with the order number of x'', replace the original word in text document , finishing watermarking embedding with 1-bit.

6) To traverse the whole text of the sentence.

7) Connecting the binary of overflow and underflow information into string L , according to the chaotic search algorithm, recording L in chaos sequence as a sequence, and saving the initial value y_0 and the starting position B as secret key, then finishing embedding algorithm.

B. Extraction Algorithm

The processing of extraction algorithm is the inverse processing of embedding algorithm. The implementation steps as follows:

1) With step 1 and step 2 of the watermark embedding algorithm, finding the initial value x'' , and the predicted value $\hat{x}$.

2) Calculating $p_e'' = x'' - \hat{x}$,extracting watermark information. According to (2), the watermark w is related with $LSB(p_e')$. If $LSB(p_e') = 1$, $w = 1$; If $LSB(p_e') = 0$, $w = 0$. Therefore, according to ***Rule 2***, we can get watermark information from calculating $p_e'' = x'' - \hat{x}$. Divide into three conditions:

a) if x'' belongs to Class I, calculate $p_e'' = x'' - \hat{x} = x' - \hat{x} = p_e'$, get $LSB(p_e'') = LSB(p_e')$;

b) if x'' belongs to Class II, calculate $p_e'' = x'' - \hat{x} = -x' - \hat{x}$, and $p_e' = x' - \hat{x}$, get $LSB(p_e'') = LSB(p_e')$;

c) if x'' belongs to Class III, calculate $p_e'' = x'' - \hat{x} = \hat{x} - 1 - \hat{x} = -1$, and $p_e' = x' - \hat{x} = \hat{x} + 1 - \hat{x} = 1$, get $LSB(p_e'') = LSB(p_e')$.

Therefore, $LSB(p_e'') = LSB(p_e')$. So, if $LSB(p_e'') = 1$, watermark information $w = 1$; If $LSB(p_e'') = 0$, watermark information $w = 0$.

3) To traverse the whole text, finishing extraction algorithm.

C. Reversible Algorithm

Through the reversible algorithm, the watermarked text will be reverted to the original text.

1) With step 1 and step 2 of the watermark embedding algorithm, finding the initial value x'' , and the predicted value $\hat{x}$.

2) Using the secret key y_0 and B , through chaotic sequence to create the overflow and underflow information. According to ***Rule 1***, get the classification information of every position.

3) According to the corresponding classification information, through the ***Rule 2***, restore out of every position x' .

4) According to Eq. 2, calculating to get the initial value x of original words, and recovering.

5) To traverse the whole text, finish reversible algorithm.

4. Experimental Results and Analysis

Since proposed algorithm through the CCD to choose word after word segmentation, uses the prediction-error expansion algorithm to realize watermark embedding, and stores the overflow and underflow message in chaos sequences. Therefore, according to the principle of the prediction-error expansion algorithm, we can get the watermark information without any other information in extraction algorithm. In reversible algorithm, through the initial value and the starting position of chaotic sequence, calculating the overflow and underflow message, using the reverse prediction-error expansion algorithm can restore the original text expansion intact.

The following parts are the experimental simulation results. The original text document is shown in Fig. 3, the shaded parts are the located words. Assuming that the embedding watermark sequence is 100111001, the threshold $T=0$ .Through the simulation, invisibility, robustness, reversibility, security and contrast analysis of algorithm are discussed respectively.

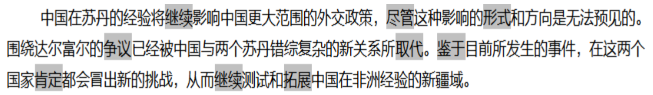

中国在苏丹的经验将继续影响中国更大范围的外交政策，尽管这种影响的形式和方向是无法预见的。围绕达尔富尔的争议已经被中国与两个苏丹错综复杂的新关系所取代。鉴于目前所发生的事件，在这两个国家肯定都会冒出新的挑战，从而继续测试和拓展中国在非洲经验的新疆域。

Fig. 3 Original text

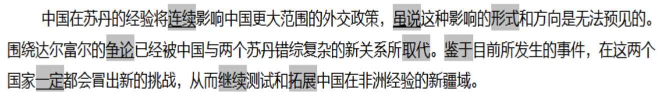

中国在苏丹的经验将连续影响中国更大范围的外交政策，虽说这种影响的形式和方向是无法预见的。围绕达尔富尔的争论已经被中国与两个苏丹错综复杂的新关系所取代。鉴于目前所发生的事件，在这两个国家一定都会冒出新的挑战，从而继续测试和拓展中国在非洲经验的新疆域。

Fig. 4 Watermarked text

A. Invisibility

The watermarked text is shown in Fig. 4, and the words changed are marked with single underline. Compared Fig. 3 and Fig. 4, we can see that, through the replacement of synonym, the watermarked text don't cause any format appearance change; Using the CCD algorithm avoid semantic ambiguity efficiently. This algorithm has good invisibility.

B. Robustness

The extracting watermark information after text regeneration attack, text format conversion, font adjustment, deleting spaces, and punctuation unified revisions are given respectively in Table 1. Text regeneration attack is rewriting the original text; Text format conversion is the text font transformation, just as converting WORD into TXT; Font adjustment is the word font transformation, such as song typeface into FangSongTi; Deleting spaces is removing the excess spaces; Punctuation unified revisions is the exchange of Chinese and English punctuation. From Table I, we can see that these format attacks don't break the watermark information, because the algorithm uses the natural language processing technology to embed watermark and the embedding process occurs in the content of text, and it doesn't destroy watermark information, while meeting with format attack, this algorithm has strong robustness.

Table I Format Attack Analysis

The kinds of attack	The extracting watermark information
Text regeneration attack	100111001
Text format conversion	100111001
Font adjustment	100111001
Deleting Spaces	100111001
Punctuation unified revisions	100111001

C. Reversibility

Replacing back the original word which had been replaced in embedding algorithm is called reversibility. The reversibility means the embedded reversible. The reversible algorithm not only requires extracting the embedded watermark, but also requires recovering the words replaced in original text, realizing the reversibility of the original text content. The recovered text is given in Fig. 5. Compared Fig. 3 and Fig. 5, we can see that the algorithm can realize the watermarked text nondestructive recovery.

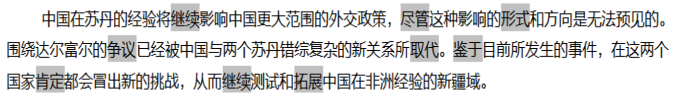

中国在苏丹的经验将继续影响中国更大范围的外交政策，尽管这种影响的形式和方向是无法预见的。围绕达尔富尔的争议已经被中国与两个苏丹错综复杂的新关系所取代。鉴于目前所发生的事件，在这两个国家肯定都会冒出新的挑战，从而继续测试和拓展中国在非洲经验的新疆域。

Fig. 5 Recovered text

D. Security

The security of traditional algorithm by synonyms replacement is guaranteed by the uncertainty of synonym library, and the synonym library is usually kept secret. In this paper, the synonym library comes from HIT IR Lab [12], which is open. Therefore, the security in the algorithm is to say the attacker can't get the original text, in the condition of unknown the initial value and the starting position information in chaos sequence. Fig. 6 gives the recovered text for the initial value unchanged, but the starting position information changed; Fig. 7 gives the recovered text for the initial value changed and the starting position information unchanged. Fig. 8 gives the recovered text for the initial value and the start position information all changed, the words have not accurate recovered marked with double underline. From Fig. 6 to Fig. 8, we can see, the recovery results have bigger difference with the content of the original text. That is to say, if the attackers haven't get both the initial value and the starting position information, he can not accurate recovering the original text, the algorithm has strong security.

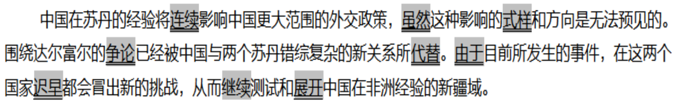

中国在苏丹的经验将连续影响中国更大范围的外交政策，虽然这种影响的式样和方向是无法预见的。围绕达尔富尔的争论已经被中国与两个苏丹错综复杂的新关系所代替。由于目前所发生的事件，在这两个国家迟早都会冒出新的挑战，从而继续测试和展开中国在非洲经验的新疆域。

Fig. 6 Recovered text for initial value not change and start position change

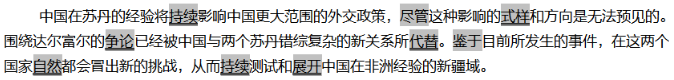
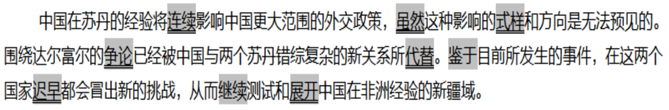

中国在苏丹的经验将持续影响中国更大范围的外交政策，尽管这种影响的式样和方向是无法预见的。围绕达尔富尔的争论已经被中国与两个苏丹错综复杂的新关系所代替。鉴于目前所发生的事件，在这两个国家自然都会冒出新的挑战，从而持续测试和展开中国在非洲经验的新疆域。

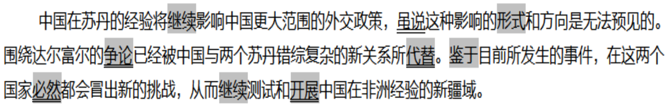
Fig. 7 Recovered text for initial value change and start position not change



中国在苏丹的经验将连续影响中国更大范围的外交政策，虽然这种影响的式样和方向是无法预见的。围绕达尔富尔的争论已经被中国与两个苏丹错综复杂的新关系所代替。鉴于目前所发生的事件，在这两个国家迟早都会冒出新的挑战，从而继续测试和展开中国在非洲经验的新疆域。

Fig. 8 Recovered text for initial value and start position change all change

In addition, if the threshold value T changes, the extracting watermark changes and the result of recovering also changes. Fig. 9 gives the result of recovering when the threshold value $T = 10$, and the extraction of watermark is 10111001. We can see from Fig. 9, the location of the words have changed; The recovering of the content has bigger difference from the original text. Therefore, the choice of the threshold value T can increase the safety of this algorithm further.



中国在苏丹的经验将继续影响中国更大范围的外交政策，虽说这种影响的形式和方向是无法预见的。围绕达尔富尔的争论已经被中国与两个苏丹错综复杂的新关系所代替。鉴于目前所发生的事件，在这两个国家必然都会冒出新的挑战，从而继续测试和开展中国在非洲经验的新疆域。

Fig. 9 Recovered text for threshold T change

E. Experimental analysis compared to other algorithms

Experiments also give the comparison with other reversible text watermarking algorithm, shown in Table II. Through the comparison, we can see that, compared with the method in [7], the algorithm hasn't caused semantic ambiguity after synonym substitution, increasing the invisibility of the algorithm; And word library is open, the secret key can keep the security of the algorithm better. Compared with the method in [8], the algorithm avoids a lot of auxiliary information to restore text, it is easy to practical application.

Table II Comparison of Related Algorithms

The system of watermark	Method in [7]	Method in [8]	Our Method
Whether caused deviation in the semantic	Yes	no	no
Defense of format attack	Strong	Strong	Strong
Defense of synonym replacement attack	Weak	Weak	Weak
Security(Synonym word library open)	Weak	Strong	Strong
Append informatin in reversible algorithm	Nothing	Much(need lots of XO information)	Little(the secret key)

5. Conclusion

In this paper, through the calculation of the context collocation degree, locating the embedding watermark position, eliminating the semantic ambiguity after words replacement, with the chaotic sequence to determine the overflow and underflow information, translating safety to chaos mapping with selecting the initial value and ensuring the starting position, discusses a reversible text watermarking algorithm using prediction-error expansion. According to the located watermark position with the context collocation degree, we can get the watermark information using the prediction-error expansion algorithm, without any other information in extraction algorithm. Using the secret key, the algorithm can restore the original text expansion intact. Simulation experiments show that the algorithm not only satisfies good invisibility, but also shows a strong robustness in the format attack. This algorithm has strong security in extraction and reversible algorithm. But the algorithm also has certain disadvantages, such as it can't resist the attack in content, which will be solved in the future research.

6. References

[1] Feng J B, Lin I C, Tsai C S, et al, Reversible watermarking: current status and key issues[J]. International Journal of Network Security, 2006, vol. 2, no. 3, pp. 161-171.

[2] Tian J. Reversible data embedding using a difference expansion[J]. IEEE Trans on Circuits and Systems for Video Technology, 2003, vol.13, no. 8, pp.890-896.

[3] Thodi D M, Rodriguez J J. Reversible Watermarking by Prediction-Error Expansion[C]. Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation, on Image Analysis and Interpretation, 2004, vol. 5, pp. 21-25.

[4] Hu Yong-jian, Lee Heung-Kyu, Li Jian-wei. DE-Based Reversible Data Hiding With Improved overflow and underflow Location Map [J]・IEEE Transactions on Circuits and Systems for Video Technology, 2009, vol.19, no. 2, pp. 250-260.

[5] Chrysochos E, Varsaki E E, Fotopoulos V, et al.High capacity reversible data hiding using overlapping difference expansion[C]. Proc of Image Analysis for Multimedia Interactive Services,London,May. 6-8, 2009, pp. 121-124.

[6] Coltuc, Dinu. Improved embedding for prediction-based reversible watermarking[J]. IEEE Transactions on Information Forensics and Security, September 2011, vol. 6, pp. 873-882.

[7] Liu zhi-jie. Research on Reversible Text Watermarking Based on Natural Language[D]. [Master Thesis]. Hunan: Hunan University, 2010.

[8] Jiang Chuan-xian, Chen Xiao-wei. Robust Reversible Text Watermarking Algorithm[J]. Journal of Computer-Aided Design & Computer Graphics, 2010, vol. 22, no. 5, pp. 879-885.

[9] Topkara M, Topkara M, Mercan, Atallah M J. The hiding virtues of ambiguity: Quantifiably resilient watermarking of natural language text through synonym substitutions[C]. Geneva: Proceedings of ACM Multimedia and Security Workshop(MMSEC'07), 2006, pp. 164-174.

[10] Gan can, Sun Xing-ming, Liu Yu-ling, Xiang Ling-yun. Animproved Steganograghic Algorithm based on Synonymy Substitution for Chinese Text[J]. Journal of Southeast University in China, 2007, 31, Sup(I), pp.137-140.

[11]Zheng Xue-ling, Huang Liusheng, Chen Zhili,et al. Hiding Information by Context-Based Synonym Substitution[C]. Digital Watermarking - 8th International Workshop Proceedings, Guildford, United kingdom, August 26, 2009, pp.162-168.

[12] Wanxiang Che, Zhenghua Li, Ting Liu. LTP: A Chinese Language Technology Platform. In Proceedings of the Coling 2010:Demonstrations. 2010.08, pp.13-16, Beijing, China.

[13] Sougou Labs. SougouR. http://www.sogou.com/labs/dl/r.html, 2011

[14] Xiang Hua, Cao Han-qiang, Wu Kai-ning, Wei Fang. A Zero-watermarking Algorithm Based on Chaotic Modulation[J]. Journal of Image and Graphics, 2006, vol.11, no.5, pp. 720-724.