

# A Method of Semantic Hidden Reduction Based on Collocation

Licai Zhu<sup>1</sup>

<sup>1</sup> Yancheng Teachers University, Yancheng, Jiangsu, China

Correspondence: Licai Zhu, Yancheng Teachers University, Yancheng, Jiangsu, China.

Received: September 20, 2017

Accepted: October 30, 2017

Online Published: October 31, 2017

doi:10.5539/cis.v10n4p73

URL: <http://doi.org/10.5539/cis.v10n4p73>

## Abstract

Semantic hiding is the technology of using semantic knowledge to embed secret information into text carrier. Among the many methods of semantic hiding, "synonym substitution" is paid more and more attention by semantic hiding. The main idea of this method is to hide the secret information by replacing synonyms in text so as to retain its original meaning as much as possible. In order to effectively restore hidden information, we need to find the synonym replacement location as accurately as possible, so it is very important to recognize the collocation of words. So far, however, there is no effective way to identify and match Natural Language Processing, that is, it is very difficult to tell exactly whether or not the words in the text have been replaced.

In this paper, a hidden reduction method based on collocation is proposed. By analyzing the characteristics of synonyms and their collocation, this paper treats their relation as the relation between the pairs of samples in statistical sense. According to the nature of the statistic, we design several decision features to identify the collocations. At the same time, we introduce the form of point mutual information in the information theory as a feature to use the independence of quantifier pairs. In order to recognize word collocation effectively, this paper combines these features, and uses genetic algorithm to get the recognition degree of each feature. Then, a replacement recognition system based on immune abnormality mechanism is designed. Synonyms for collocation are regarded as "normal", while substitutions are regarded as "anomalies". The experimental samples are generated by semantic hidden software TLEX. To better render the restore process, we rewrote the TLEX to add the key selection module.

**Keywords:** Hidden reduction, semantic concealment, word collocation, synonym substitution, TLEX, statistical features

## 1. Introduction

With the widespread use of text media, more and more attention has been paid to the technology of using text as a carrier for information hiding (Anderson & Kuhn, 1999). Semantic hiding is an important branch of the field of text hiding. It mainly uses semantic knowledge to embed secret information into text vectors. When hiding information, semantic hiding technology preserves the original meaning of the carrier as much as possible, and hides the secret information very well, which makes it difficult to find the modified location of the dense text. Therefore, semantic hiding technology meets the needs of information hiding applications, and greatly enhances the security of information transmission and storage.

In the field of semantic hiding, it is a good choice to make use of the similar meanings between synonyms to hide information. For example, there is an English sentence "There is a big room". If the word "big" is replaced by "large", it is obvious that the replaced sentence will not affect the meaning and acceptability of the original sentence. In other words, it has almost the same meaning as the original sentence, so the replaced sentence is semantically acceptable. In addition, there are a number of synonyms, such as "propensity", "predilection", "penchant" and "proclivity", there are multiple substitutions may. The synonym substitution method maximizes the preservation of the source text without changing the text of the enclosed text in terms of morphology, syntax, and semantics. Compared with other methods such as hidden Markov chain or semantics, given grammar rules, this method has robustness and adaptability, more so by the synonym substitution method of hidden information is difficult to be detected and reduction.

Among the many current text hiding tools, TLEX (Winstein, 1999) is the famous semantic hiding tool which uses synonym substitution idea to hide information. This paper is to study the information reduction of dense text generated by TLEX. This paper refers to the public key cryptography (Diffie & Hellman, 1976). It is

assumed that the hidden information in the text has been known before the reduction, and the secret method is known. Therefore, the security of the hidden tool depends on the choice of key to the hidden bits. In accordance with this principle, this article rewrites TLEX to add a hidden bit selection module. The module uses the random location generated by the key mechanism to determine where the synonyms can be replaced. In order to restore secret information, our task is to find the hidden bits as accurately as possible to predict the key. As far as the author is concerned, since Cuneyt, M., Taskiran and other (Taskiran, et al., 2006) people studied textual information hiding detection for the first time, this paper is the first to discuss and study the problem of text information hiding restoration.

This paper proposes a semantic recovery method based on collocation. The method is based on the fact that the use of the word is limited to polysemy, background, use domain and common collocation constraints, although synonyms can be replaced to retain the source text meaning, but may destroy the collocation constraint of these words. By analyzing synonyms and collocation constraints, the relation between them can be regarded as the relation between pairs of samples in statistical sense. In this sense, this paper identifies the collocation of words according to the statistical design of different discriminant features. In order to further improve the recognition rate, this paper considers combining these features. We first use genetic algorithm to obtain the recognition degree of each feature, and then build a replacement recognition system based on immune anomaly mechanism. Trained to generate a set of valid detectors that replace the synonyms in the hidden text as an exception attack.

The method is outlined as follows: first, define a collocation window, N-WINDOW, which matches each word pair in the window as a candidate pair. Get all the words around the synonyms in the natural text with N-WINDOW; Then, the statistical knowledge is introduced, and the five words are set to judge the features. To detect the recognition rate of each feature, a genetic algorithm is used to evaluate them. According to the recognition rate of feature recognition, this paper designs an alternative recognition system based on immune anomaly mechanism. It replaces synonyms for the destruction of collocation as "abnormal". The system generates a set of detectors to detect these exceptions. When the information is restored, the system identifies the replacement bits in the hidden text, and then predicts some of the keys.

The following is organized as follows: in the second section, the basic idea of synonymous substitution is given, and the process of TLEX hiding information and the rewriting of the software are briefly described; The third section gives the characteristics of collocation judgment; The fourth section uses genetic algorithm to evaluate the recognition of each feature to construct the collocation evaluation system, and then restore the information; In section fifth, experiments are conducted to evaluate the performance of the proposed method. Finally, the sixth section is the conclusion.

## 2. Synonyms Replacement Method

### 1) Basic Ideas

The basic idea of the synonyms substitution method is that the hidden person finds the synonym S that can replace the original word E, and then uses the hidden function  $s(E, w) = S$  to hide the secret information W. After receiving the hidden information, the receiver decrypts the secret text by using the decryption function  $d(S) = w$ .

Therefore, the selection of alternative words is the key to successful concealment. Only words that can preserve the meaning of the source text can be considered. Set the degree of freedom of words D (degree, of, freedom), which is determined by the meaning of the word. In the selection process, if DS is less than or equal to DE, it can be replaced by S E, and vice versa. According to degrees of freedom, when you describe basement, you can replace dank with clammy; you can use praiseworthy instead of laudable to describe accomplishment; gallant knight can also be replaced by chivalrous knight.

Generally speaking, the hidden text generated by the synonyms substitution method has two ways of decryption, that is, source text and text are not needed. For the first method, the source text is compared with the stored text to decrypt the changes in the latter to restore the hidden information; For the second approach, you use the pre-set replacement rules to find the places in the text that are loaded to restore hidden information. TLEX belongs to the latter way.

### 2) TLEX Practices

TLEX is a text hiding tool that uses the synonyms substitution method to hide information. The tool first gets the meaning of each word from the WordNet (Beckwith & Miller, 1990) and calculates their degrees of freedom. Then, the words with the same degree of freedom are formed into synonyms, and then the thesaurus data can be used to replace them. To hide the secret information, TLEX finds all synonyms that can be used in the natural text according to data, and then uses these synonyms as indexes to sort and number the synonyms for the index

words in data. For example, data has synonyms for array={unremarkably, commonly, normally, ordinarily, usually}, and assumes that the vector text contains "we, usually, sleep, the, siesta, after, our, lunch". According to data, TLEX to find the "usually" is a synonym replacement, then array will sort new synonym of array '= {commonly, normally, ordinarily, unremarkably, usually}, and array elements' number {commonly (0), normally (1), ordinarily (2), unremarkably (3) usually (4)}. According to the secret information, TLEX selects the corresponding synonym to hide.

### 3) overwrite TLEX

In order to show the hidden restore process, with reference to the public key cryptography, this paper assumes that the hidden information in the text containing the known hidden form has been determined before the restoration, and the security of the hidden tool relies on the choice of key. Based on this principle, this article rewrites TLEX, adding the key selection to hide the location of the module. When information is hidden, TLEX determines whether the selected bits can be selected according to the given key. For the rewritten TLEX, the selection of the hidden bits is determined by the key and thesaurus. Therefore, in order to restore secret information, it is necessary to find the secret location prediction key accurately.

Concretely, the working procedure of the key bit selecting module is that the key selecting module is:

(1) setting a N bit key

(2) using this key as the seed of the random number generator, a series of random numbers R1, R2 are generated,... RN.

(3) each bit value of  $\text{posi} = \text{ri} \% 2$  to be hidden. If  $\text{posi} = 1$ , the bit is allowed to hide, otherwise it cannot be hidden.

Generally speaking, the parity ratio of random numbers is equal, that is to say, the length of the random number to be generated is  $n = 2 * \text{the secret information becomes the length of the binary code}$ .

### 3 Judgment Feature of Collocation

Obviously, if the words in the text have a certain inherent collocation, then replace it with synonyms, the original collocation will be destroyed. In view of this fact, this paper considers the collocation of identifying words to determine the hidden bits in dense text.

In Natural Language Processing, word collocations are descriptions of the idiomatic positions of the word, (Firth, 1957) including noun phrases, verb phrases, and fixed phrases. The standard measure of collocation in typical linguistics is (Benson, 1989; Brundage et al., 1992), which consists of non-constituent word formation (Non-compositionality), Non-substitutability, and Non-Modifiability. For example, strong and powerful can be regarded as a pair of words of similar meaning. But in English, strong and tea are idiomatic collocations, while powerful tea is not, that is, the latter is a habitual expression.

The word collocations defined in this paper include not only the fixed collocation under the typical definition, but also the phenomena of Association (Association) and co-occurrence (co-occurrence). This is because some words are not typical, but they are strongly related to each other, such as plane-airport. Increasing the recognition of these connectives can be more effective in judging synonymous substitutions.

In order to determine whether the words match or match, this paper designs different judgment characteristics. These features can be used for collocation recognition, and in many cases the ability to identify is roughly the same, but also has its own range. The decision features are designed as follows:

#### 1) Variance

Since collocation is a local phenomenon, this paper sets up a collocation window "N-WINDOW". Explore the relationship between synonyms W and its surrounding N words. Specific approach is: If there is a sentence in the text,  $S = \dots w_{i-1} \dots W w_{i+1} \dots W \dots$ ,  $w_{i-1}$  - W and W distance were  $D(W_{i-1}, w)$ ,  $\dots$ ,  $D(W, w) = \dots$ .  $D(W_i, w)$  is considered as the offset of  $w_i$  relative to W. Based on the statistics of natural text all meet - less than  $D(W_i, w)$   $w_i$  and W is less than or equal to the offset,  $i=1, \dots, N$ . Calculate the variance of the offset of  $W_i$  and W s. According to statistical knowledge, the smaller the variance of the offset of the two things, the higher the degree of intimacy between them. If the variance of W and  $w_i$  is very small, they can be considered as collocation.

However, some low variance words obtained by this method may occur by chance, and then they cannot form a match. Therefore, we consider the use of hypothesis testing to evaluate. Define the null hypothesis (null, hypothesis)  $H_0: W$  and  $w_i$  appear independently. Under the hypothesis, calculate the probability of occurrence of the event P, if the  $p >$  confidence level (generally less than 0.005), be sure  $H_0$ , otherwise, reject  $H_0$ . The event probability is obtained by t test and Pearson chi square test.

## 2) t test

Under the condition that the expected mean and the observed mean are different, the t test can assume that the sample obeys the normal distribution of the mean, and it is concluded that the likelihood of the test data to satisfy the same mean and variance is large.

Statistics of T detection

$$t = \frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{N}}} \quad (1)$$

Among them is the sample mean, the distribution mean,  $S^2$  is the sample variance, and  $N$  is the sample size.

In this paper, t detection is introduced into collocation recognition. Firstly, the natural text set is regarded as a sequence set consisting of  $L$  two word pairs, and the synonym pairs are considered as data samples. If the word "W1" appears  $T_1$  times, W2 appears  $T_2$  times, and  $w_1w_2$  (or  $w_2w_1$ ) appears  $T_3$  times. The null hypothesis is that  $H_0: w_1$  and  $W_2$  are independent. The conditions are obtained by  $=P(W_1)$ ,  $P(W_2)$ ,  $=P(w_1w_2)$ ,  $s_2 = P(w_1w_2) * (1 - P(w_1w_2))$ ,  $N=L$ . According to (1) calculate the statistic  $T$  and compare with the critical value  $t'$  of  $=0.005$ , if  $t > t'$ , then reject the null hypothesis:  $W_1$  and  $W_2$  do not appear independently, they can form a match.

By using the judgment feature, we can judge whether the word pairs exist or not.

3) chi square test of Pearson( $\chi^2$ )

Although the t test can be used to identify collocations, the drawback of this test is that the prior distribution of the data is assumed to be normal distribution prior to detection. However, this assumption is not always met in all cases. For this reason, another method of inspection, Pearson chi square test, is introduced to judge the collocation. Unlike t testing, it does not require data samples to satisfy normal distribution.

According to the statistical knowledge, the sum of the difference between the observed and expected values is calculated, and the expected value is used as the scaling factor to obtain the difference

$$\chi^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (2)$$

Among them,  $I$  and  $j$  represent samples, and  $O_{ij}$  represents the observations of the samples, and  $E_{ij}$  represents the expected values.

This article is concerned with the relation of two - yuan words to ( $W_1$ ,  $W_2$ ), so that statistics become

$$\chi^2 = \frac{N(O_{11}O_{12} - O_{21}O_{22})^2}{(O_{11}+O_{12})(O_{11}+O_{21})(O_{12}+O_{22})(O_{21}+O_{22})} \quad (3)$$

Among them,  $O_{11}$ ,  $O_{12}$ ,  $O_{21}$  and  $O_{22}$  represent  $w_1w_2$ ,  $w_1w_2$ ,  $w_1w_2$  and  $w_1w_2$  (denoted non), and  $N$  is the total number of words pairs.

Through the statistics of the two words in the corpus, we get the parameter value, and then calculate the statistics according to (3). The value is then compared with the critical value of the statistics at  $=0.005$ . If the former is large, it means that  $w_1w_2$  can form a match, otherwise it means independence.

Compared with the t test, which is more suitable for Pearson chi square detection method in probability value is relatively large, but Snedecor and Cochran (Snedecor et al., 1989) pointed out in the statistical sense, if the sample size is less than 20 or even between 20-40 but the expected value of the statistic is less than 5, the detection is not suitable to be used. Therefore, if the word is not large enough for the sample, the statistic is set to 0.

## 4) likelihood ratio

As mentioned above, the t test can detect the normal distribution of the words to the sample, and the Pearson chi square test has a better detection effect on the larger sample pairs. However, for collocation, more States belong to sparse collocation. In order to solve this problem, the likelihood ratio is used to express the relation of words.

For the two element word, for  $w_1w_2$ , if  $P(w_2|w_1) = p_1$ ,  $P(w_2|w_1) = p_2$ , give two hypotheses about the two tuple: Suppose  $1:p_1 = P_2$

If  $2:p_1 = P_2$

Here, suppose that 1 is the independent assumption of  $W_1$  and  $W_2$ , and assume that 2 is a non-independence hypothesis. Similar to the t test, the word " $W_1$ " appears  $T_1$  times,  $W_2$  appears  $T_2$  times, and  $w_1w_2$  (or  $w_2w_1$ ) appears  $T_3$  times. The total number of two words in the corpus is  $L$ ,  $p_1 = t_2/t_1$ ,  $p_2 = (t_2 - t_1) / (N - t_1)$ . When  $p_1 = p_2$ , the value of the likelihood ratio  $\lambda$  is also represented by  $p = c_2/N$

$$\begin{aligned} \log \lambda &= \log L(H_1) / L(H_2) \\ &= \log L(c_{12}, c_1, p) + \log L(c_2 - c_{12}, N - c_1, p) \\ &\quad - \log L(c_{12}, c_1, p_1) - \log L(c_2 - c_{12}, N - c_1, p_2), \end{aligned} \quad (4)$$

Among them,  $L(k, N, x) = \binom{N}{k} x^k (1-x)^{n-k}$ .

According to Mood and other conclusions in the literature (Mood, et al., 2000), the  $-2\log$  approximation is asymptotically approximated in statistics, so the expression is used when deciding the two tuple relationship. After calculation, if the value of  $-2\log \lambda$  is greater than the critical value of the confidence level  $= 0.005$ , then  $H_1$  is rejected, or  $H_2$  is rejected.

When the two word pairs belong to the sparse collocation, the likelihood ratio is used as the judgment feature, which has better recognition effect than the first two statistics. However, when the word pair is too sparse (the expectation value is less than 1), the approximation effect of  $-2\log \lambda$  is not very good, (Pedersen, 1996), in other words, its recognition effect will become very poor.

#### 5) mutual information (MI)

Mutual information (MI) is the amount of information associated with two events, (Fano, 1961), which is used in information theory to represent the connection between random variables. This article uses MI to identify words and phrases.

The mutual information of the two tuple  $w_1w_2$  is as follows:

$$I(W_1, W_2) = \log_2 P(w_1w_2) / P(W_1) P(W_2) \quad (5)$$

By (5) we can see that when  $W_1$  and  $W_2$  are completely independent, there is:

$$\begin{aligned} I(w_1, w_2) &= \log_2 P(w_1w_2) / P(w_1)P(w_2) \\ &= \log_2 P(w_1)P(w_2) / P(w_1)P(w_2) \\ &= \log_2 1 \\ &= 0 \end{aligned}$$

This shows that mutual information can well identify independent word pairs. However, it is difficult to identify dependent words. To illustrate this, consider the extreme dependence of the word on the  $w_1w_2$  interdependence - at this point

$$\begin{aligned} I(w_1, w_2) &= \log_2 P(w_1w_2) / P(w_1)P(w_2) \\ &= \log_2 P(w_1) / P(w_1)P(w_2) \\ &= \log_2 1 / P(w_2) \end{aligned}$$

From the upper model, we can see that for the fully dependent word pairs, their mutual information will decrease when the number of them appears increases.

According to the above analysis, the four characteristics of this design can be used for feature recognition, but the scope of application is different. When judging collocations, the results of different features may vary, and sometimes the difference may be great. Therefore, the evaluation of each feature's recognition capability is critical.

## 4. Building Alternative Identification Systems

### 4.1 Feature Recognition

In this paper, the recognition ability is used to evaluate the recognition ability of feature pairs. Obviously, the recognition degrees of each decision feature are different. In order to recognize the collocations effectively, it is necessary to evaluate their recognition accuracy.

In order to evaluate the degree of recognition, the genetic algorithm is used to obtain the degree of feature recognition. In this paper, the fitness function required by the genetic algorithm is  $f(x)$ , which represents the feature recognition degree and represents the decision value of the matching degree of the feature pairs. After

training, the recognition of each feature is obtained. Calculation of training samples in D on the collocation of degree:

$$D = \begin{cases} D_s, \text{ if } D_s \leq \varepsilon, \\ D_i, \text{ if } D_i \geq D_j, D_j \in \{\dots D_{i-1}, D_{i+1}, \dots\} \\ \frac{\sum_{i=1}^5 D_s}{5} \end{cases}$$

#### 4.2 Replacement Identification System

If will replace the damage caused by hidden in collocation synonyms as attacks, then, can be regarded as incorrect collocation of "anomaly objector", the correct collocation as "normal objector". According to this idea, this paper uses the immune anomaly mechanism to construct the replacement identification system.

The mechanism of immune abnormality is proposed by Forrest, which is mainly used in computer intrusion detection (Forrest & Hofmeyr, 2000). So far, the technology has developed to a great extent and has been used effectively in many fields (Kwon & Nasrabadi, 2005; Steinwart et al., 2006; Xiang & Gong, 2008). The anomaly detection in this paper is an anomaly collocation analysis process. It receives the N-WINDOW pairs of the hidden text synonyms, while the abnormal pairs are identified by the non-set pairs. In the process of recognition, if the word pairs belong to the nonself set, then the word pairs are found to be abnormal. Then, the system calculates the anomaly probability according to the feature recognition degree.

If there is a synonym phrase "array={w1", "W2",... Wn}", the collocation set of word wi is {wi-wi1, wi-wi2,... Wi-wim}, where m=|wi|, wi-wij stands for {wiwij or wijwi}, j=1,... M. The abnormal immune mechanism, self set (Self) = {wi, wi-wij | wi in array}, the initial Noe-Self set wa-wbk | = {nonself set a = B, 1 = a, B = n, k = 1 ~ |wb|}.

A group of effective detectors are obtained after training. In the process of detection, each detector predicts the degree of word substitution.

#### 4.3 Restore Information

According to the substitution recognition system, the synonym substitution degree is obtained. This article divides the substitution degree into two kinds:

Class 1: replacement degree < setting threshold dt,

The second category: threshold dt < substitution degree < 0.5.

When the information is restored, the part of the key is determined according to the collocation of the first class. If the key number is less than N-k, k = 5 ~ 10, according to second kinds of situations, choose some position as a replacement to predict key. For the last I bit key, the exhaustive method is used for analysis.

### 5. Experiment and Analysis

#### 5.1 Experiment

Experimental samples: rewritten TLEX generation

Key length: N=20

Hidden information length: 100b, 200B, 500B

Restore the result of 100b

Non predicted digit	Reduction time(ms)	Accuracy rate
10	955	87.8%
9	584	91.6%
8	298	94.1%
7	214	97.5%
6	135	98.9%
5	102	99.3%

Reduction time = system evaluation time + exhaustive time, in which the exhaustive time is divided into two

kinds: If the prediction of the previous N-i bit key is accurate, the exhaustive time is the time of the full permutation of the I bit key; Otherwise, the exhaustive time also contains the prediction time. Accuracy refers to the accuracy of the information reduction, regardless of whether the former N-i bit key is accurate or not.

### 5.2 Performance Analysis

As stated above, the performance of the present method depends on the exact number of digits and the time to restore. It is apparent that the reduction time is largely influenced by the accuracy of the first prediction. Here are the best, worst, and average restores times.

Since the time used in the evaluation phase has not changed much, the unified setting is E. Let  $T_k$  indicate the time needed to verify the k bit key using exhaustive method.

Best case: N-k sites were accurately predicted with a reduction time of  $=E+T_k$

Worst case: N-k are the error prediction, reduction time  $= E + C_{N-k}^1 * T_1 + C_{N-k}^2 * T_2 + \dots + C_{N-k}^{N-k} * T_{N-k} + (N-k) * T_{N-k}$

Average: expected restore time  $= E + C_{N-k}^1 * T_1 * p^{N-k-1} + C_{N-k}^2 * T_2 * p^{N-k-2} (1-p) + \dots + C_{N-k}^{N-k} * T_{N-k} * (1-p)^{N-k} + (N-k) * T_{N-k}$ ,

Where p represents the degree of collocation of the bit. For convenience of expression, the unified expression is expressed by P, and then converted to actual value.

## 6. Conclusion

In this paper, a hidden reduction method based on collocation is proposed. By analyzing the characteristics of synonyms and their collocation, this paper treats their relation as the relation between the pairs of samples in statistical sense. According to the nature of the statistic, we design several decision features to identify the collocations. At the same time, we introduce the form of point mutual information in the information theory as a feature to use the independence of quantifier pairs. The following work are two unfolds, which are in-depth analysis of the statistics and analysis and collocation from semantic point of view.

## Acknowledgement

This work is supported by Prospective Joint Research project of Jiangsu Province under Grant No. BY2016066-05, Teaching reform project of Jiangsu Provincial Department of Education under Grant No. 2015JSJG526.

## References

- Anderson, F. A. P. R. J., & Kuhn, M. G. (1999). Information Hiding - a Survey. *Proceedings of the IEEE*, 87(7), 1062-1078.
- Beckwith, R., & Miller, G. A. (1990). Implementing a lexical network. *International Journal of Lexicography*, 3(4), 302-312.
- Benson, B. M. (1989). The Structure of the Collocational Dictionary. *International Journal of Lexicography*, 2(1), 1-14.
- Brundage, J., Kresse, M., Schwall, U., & Storrer, A. (1992). Multiword Lexemes: A Monolingual and Contrastive Typology for NLP and MT. Technical Report 232, *Journal of IBM Institute for Knowledge Based Systems*.
- Diffie, W., & Hellman, M. (1976). New Directions in Cryptography. *IEEE Transactions on Information Theory*, 22(6), 644-654.
- Fano, L. J. (1961). *Transmission of information; a Statistical Theory of Communication*. New York: MIT Press.
- Firth, J. R. (1957). A Synopsis of Linguistic Theory 1930-1955. *Studies in linguistic analysis*, Oxford: Blackwell, 1-32.
- Forrest, S. & Hofmeyr, S. A. (2000). Architecture for An Artificial Immune System. *Evolutionary computation*, 8(4), 443-473.
- Kwon, H., & Nasrabadi, N. M. (2005). Kernel RX-Algorithm: a Nonlinear Anomaly Detector for Hyperspectral Imagery", *IEEE Transactions on Geoscience and Remote Sensing*, 43(2), 388-397.
- Mood, M., Alexander, A. G., Franklin, C. B., & Duane (n/a). *Introduction to the Theory of Statistics*. New York: McGraw-Hill, 3rd edition.
- Pedersen, Ted. (1996). Fishing for Exactness", *Proceedings of the South-Central SAS Users Group Conference*,

Austin TX.

Snedecor, G., Waddel, & Cochran, W. G. (1989). *Statistical Methods*. Ames:Iowa State University Press. 8th.

Steinwart, I., Hush, D., & Scovel, C. (2006). A Classification Framework for Anomaly Detection. *Journal of Machine Learning Research*, 6(1), 211-232.

Taskiran, C. M., Topkara, U., Topkara, M., & Delp, E. J. (2006). *Attacks on Lexical Natural Language Steganography Systems*. Proceedings of SPIE International Society for Optical Engineering, Society of Photo-Optical Instrumentation Engineers, San Jose, USA, 97-105.

Winstein, K. (1999). Lexical Steganography Through Adaptive Modulation of the Word Choice Hash. Retrieved from <http://www.imsa.edu/~keithw/tlex>

Xiang, T., & Gong, S. (2008). Video Behaviour Profiling for Anomaly Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5), 893-908.

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).