Synthesis of Secure Passwords

Tomoki Sato
Hiroaki Kikuchi
Graduate School of Engineering, Tokai University,
4-1-1 Kitakaname, Hiratsuka, Kanagawa, 259-1292, Japan kikn@tokai.ac.jp

***Abstract*—This paper studies a synthesis of password to be easily identified and hardly forgot. A new synthesis method is proposed to construct a good passwords that satisfy both requirements. Our method focuses on the change of frequency of combined words. Each of two words has a high term frequency but the combination is not quite common and then the frequency of the combined words can give strong impression in our memory. In order to verify our hypothesis, the paper presents a formal definitions of impression I, conflict C and accuracy A for synthesized words. The impression is a measure based on subjective evaluation for words. The conflict represents a degree how much reduction in frequency is given by combination of two words. The degree of conflict can be evaluated by a fraction of synthesized words in a particular corpus. The experimental results shows that the above hypothesis holds with positive correlation between accuracy for memory and the impression given from the synthesized words.**

***Keywords*-password, authentication, impression, long term memory**

I. Introduction

What is *good* password?

In this paper, we study fundamental requirements in good password to remember and hard to forget. A good password composes from *common* words easy to type in. A too long word or a word never used in our daily life could spoil the usability as password. However, a good password is an *extraordinary* phrase that has hardly ever used. Such sentence gives a strong impression to our memory and hence is hard to forget. The two requirements conflict each other. Namely, *common* but *not ordinary* words are hard to find.

Nishizaka et al. proposed a new password authentication scheme supposed to be used with cellular phone that restricts keys to type in but allows to input international symbols [1]. Their scheme combines the input method *T9* with property of Japanese words and automatically generate meaningless words.

However, synthesized word by machine is not always easy to remember for human. Meaningless word spoils the usability of authentication.

Therefore, in this paper, we propose a new synthesis method for the good passwords that satisfy both requirements. Our method focuses on the change of frequency of combined words. Each of two words has a high term frequency but the combination is not quite common and then the frequency of the combined words can give strong impression in our memory. Therefore, the impression could work for our ability to remember the particular word. In order to verify our hypothesis, we present formal definitions of impression I, conflict C and accuracy A for synthesized words. The impression is a measure based on subjective evaluation for words. The conflict represents a degree how much reduction in frequency is given by combination of two words. The degree of conflict can be evaluated by a fraction of synthesized words in a particular corpus. The accuracy indicates how accurate subject can remember the given synthesized words for long term. We perform a series of experiments using Google N-gram dataset [2] as a corpus. The experiment reveals a clear positive correlation between accuracy in memorizing given synthesized words and the degree of impression.

Our contributions of this work are as follows.

1) New measures to evaluate degree of impression, conflict and accuracy in memorizing,
2) New password synthesized scheme for humans easy to remember soon but hard to forget for long time,
3) Empirical study based on Google N-gram as a corpus in terms of impression, conflict and accuracy. The experimental results show a useful coincidence between impression and ability to memorize.

The rest of this paper is organized as follows. In Introduction, we describe the background in password synthesis and a idea that motivate us to this study. In Section II, we define fundamental notions, impression and accuracy. In Section III, we show the purpose of experiment and the experimental results. The three correlations are illustrated in scatter plot. In Section IV, we conclude our study.

II. Proposed Method

A. *Degree of* Conflict

Let W_1 and W_2 be frequencies of words w_1 and w_2 in a corpus. The *composition* of w_1 and w_2 is a single word defined as w_1w_2 (concatenation of two words) with frequency S. Using major search engine, we can easily estimate the frequency of given word in a set of web pages crowed in the search engine[1]

[1]In the experiment we will present shortly, we specify as "`w1w2`" so that the parser treats two words as one word.

Table I
SAMPLE OF SYNTHESIZED PASSWORD

w_1	w_2	W_1	W_2	S
privacy	festival	1.39×10^7	1.17×10^7	2
revolution	Granma	3.97×10^7	6.5×10^6	1
corn	Sir	1.04×10^7	1.18×10^8	5
accordingly	sun set	9.87×10^6	7.58×10^7	8
eventually	fight	6.89×10^7	1.23×10^7	9000
repeatedly	fill-in	1.69×10^7	3.74×10^7	6630

Definition 2.1: A *conflict* of composition $w_1 w_2$ is a degree of inconsistency of two words defined by

$$C_x = -\frac{1}{10} \log \frac{S+1}{W_1 + W_2}. \tag{1}$$

Frequencies of words vary so much and hence we apply the logarithm operation to define the degree which helps to mitigate the difference of magnitude of two words. As C increases, the composition gives stronger impression.

B. Password Synthesis

We list the requirements for good password.

1) A good password composes from *common* words easy to type in. A too long word or a word never used in our daily life could spoil the usability as password.
2) A good password is an *extraordinary* phrase that has hardly ever used. Such sentence gives a strong impression to our memory and hence is hard to forget.

However, we note that the two requirements conflict each other. Namely, *common* but *not ordinary* words are hard to find.

Hence, we synthesize the good passwords that satisfy both requirements. First, we choose top 10,000 words as our *word pool* from Google N-gram dataset [2]. Next, we classify the dataset into smaller subsets of element of sentence, e.g., noun, verb, adjective and so on. Finally, we grade pair of two words randomly chosen from categories, ("adverb" + "noun"), ("noun" + "noun"), in the degree C.

Table I shows some samples of synthesized password. (The words are translated from Japanese to English in the table.)

C. Google N-gram

The Google N-gram is a dataset extracted from actual web pages collected via a clawer.

The Japanese has no clear delimiter for distinguish words from sentence. For example in English, punctuations such as space and comma are delimiter between words. Hence, using knowledge in natural language and dictionary, sentences in web page are divided into terms, e.g. noun or verb. N-gram is a concatenation of N terms. For instance, "ABCD" of unigram has "AB", "BC" and "CD". The Google N-gram consists of N-gram of Japanese terms and often used as a fresh corpus with many broken words.

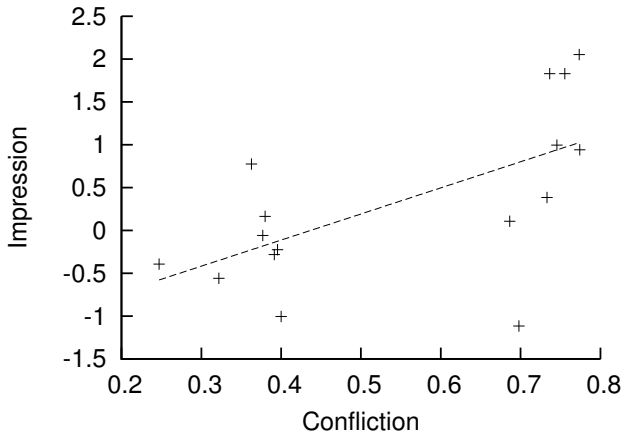

Figure 1. Impression I with respect to conflict C

III. EXPERIMENT

A. Subjective evaluation of Impression

We evaluate the degree of impression to the synthesized password.

Let $I_{x,y}$ be a degree subjective impression of j-th subject on word x, and I_x be the average for word x defined by

$$I_x = \frac{1}{n} \sum_{j=1}^{n} I_{x,j} - \overline{I_j} \tag{2}$$

where $\overline{I_j}$ is the average of all rating values evaluated by j-th subject. The rating values range from 1 (lowest) to 5 (highest impression).

B. Evaluate effectiveness to remember

In order to evaluate how accurate human can remember a synthesized words, we perform an experiment for short-term memory.

16 subjects (students) once attempt to remember four synthesized password randomly assigned from 16 synthesized words in Table I. After three days, the subjects participate to a test how accurate they can remember four words. The *accuracy* is defined as follows.

Definition 3.1: Subject j is allowed to answer password x up to three times. If the first answer is correct, let $a_j = 3$. In the first one was wrong, let a_j be 2, 1 or 0 for the second, the third correct answers or all failure, respectively. Then the *accuracy* of password x is defined as

$$A_x = \frac{1}{3n} \sum_{j} a_{j,x} \tag{3}$$

C. Experimental Result

We illustrate the relationships between two of conflict C, impression I, and accuracy A, in Figure 1, 2 and 3, respectively.

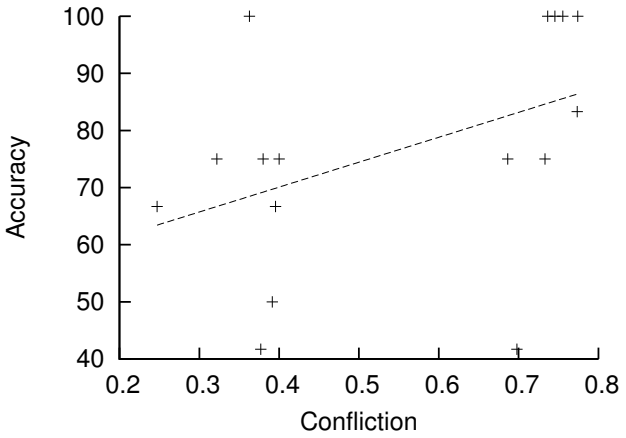

Figure 2. Accuracy A with respect to conflict C

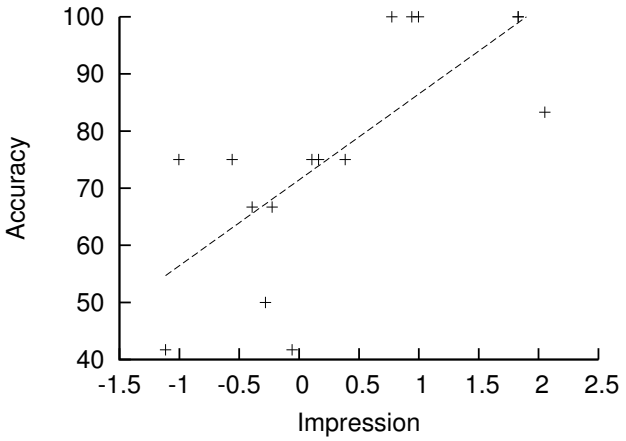

Figure 3. Accuracy A with respect to impression I

In Figure 1, a positive correlation between conflict and impression is observed. In other word, a word is likely to have higher impression as its conflict increases. The correlation coefficient is 0.617, which validate our observation is true.

Figure 2 shows that words with high conflict C perform well in terms of memory, evaluated by accuracy A. The positive correlation exists between C and A. The correlation coefficient is 0.431, which is no as strong as the relationship between C and I. There may be some factors other than impression or conflict working for our memory.

In order to clarify the reason of failure, we show Table II of password in failure to remember. The first example "privacy festival" was failed because the first word "privacy" was too strong to miss the second one. The third case was caused by two different words with same pronounce. The similar cases happen because of ambiguity of words or the way to translate into Kanji, Kana, or Hiragana forms.

There are synthesized words with high compression but low accuracy. The last example in Table II corresponds to this failure.

Table II
LIST OF PASSWORDS FAILED TO BE REMEMBERED

true	answer	reason
privacy festival	private photo	similar words
eventually funny	eventually susceptive	similar meaning[2]
first thought	begin thought	misunderstanding

In Figure 3, we observe the sharp positive correlation between impression and accuracy. The behavior is an evidence that impression is significant for remembering word in memory.

D. Security against dictionary attack

The proposed scheme is too simple and hence is vulnerable against well-known dictionary attack. In dictionary attack, adversary attempts to break a target password using an exhaustive dictionary attack with some powerful computers. Although the synthesized password from the n-gram dataset is easily to remember for human to guess, it does not immediately mean the password is secure.

To protect against dictionary attack, the password should contain non-word characters and digits so that the synthesized password cannot be constricted from knowledge of dictionary. The additional mechanism is required to add letters with the proposed password without loosing good property in terms of memory.

IV. CONCLUSION

In this study, we have proposed a new way to synthesize good passwords easy to remember. Our experiment shows a clear positive correlation between factors working in memorizing processes and the synthesized passwords perform well in term of accuracy in memory. The experiment result also reveals that uncertainty, ambiguity in language expression causes the failure in memorizing password. Our future study includes a development to prevent misunderstanding, a security against adversary who knows the algorithm of synthesize.

Acknowledge

Authors thank reviewers for careful reading and many suggestions. Authors also thank Dr. Anirban Basu for his helpful suggestion and proofreading the manuscript.

REFERENCES

[1] K. Nishizaka, M. Terada, and N. Doi, "PIN authentication using Japanese password over cellar phone", IPSJ, Tech. Report, Vol.2009-CSEC-048, pp.1-8, 2010 (in Japanese).

[2] Google Japanese n-gram data set, Japan Blog, 2007 (available at http://googlejapan.blogspot.com/2007/11/n-gram.html).