情報処理学会論文誌 Vol.56 No.9 1834–1845 (Sep. 2015)

機械合成文の不自然度相対識別問題に基づくCAPTCHAの提案

山口 通智[1,a)] 岡本 健[2,b)] 菊池 浩明[1,c)]

受付日 2014年12月5日, 採録日 2015年6月5日

概要：既存の音声型 CAPTCHA は，人間にとっても解答困難であり，視覚障害者によるウェブアクセスの障壁になっている．ワードサラダ識別型 CAPTCHA は，人間による解答容易性とアクセシビリティをあわせ持つ方式の 1 つであり，音声型 CAPTCHA の代替として期待されている．しかしながら，ワードサラダの生成に比べて自然文の収集は困難であるため，それらの問題文中での出現率の差を狙った攻撃に対して脆弱である．本稿では，自然文に相当する文もマルコフ連鎖に基づき生成することで，この問題を解決する．さらに実験を通して，文の出現率の差や検索エンジンを用いた攻撃への安全性と人間による解答容易性を評価する．

キーワード：CAPTCHA，アクセスブルデザイン，マルコフ連鎖，ワードサラダ

Proposal of CAPTCHA Based on Relative Awkwardness between Synthesized Sentences

Michitomo Yamaguchi[1,a)] Takeshi Okamoto[2,b)] Hiroaki Kikuchi[1,c)]

Received: December 5, 2014, Accepted: June 5, 2015

Abstract: Conventional CAPTCHA systems using the audio-style test are not accessible to people with visual impairments. Text-only CAPTCHA with distinguishing word salad from natural sentences is accessible and easy problems to human, is an alternative scheme. However, due to a difference of the relative frequency between word salad and natural sentence in the test, it is vulnerable against adversaries who use search engines. In this paper, we address the problem using machine-synthesized sentences by Markov chain as natural sentences. We evaluate the security against attacks with search engines. We show the experimental results of our proposal regarding usability.

Keywords: CAPTCHA, Accessible design, Markov chain, Word salad

1. はじめに

1.1 背景と目的

CAPTCHA（Completely Automated Public Turing test to tell Computers and Humans Apart）[19] は，人間とロボット（ソフトウェアによる自動化エージェント）を識別するテストである．CAPTCHA は，オンラインサービスにおいて，ロボットによる不正利用目的のアカウント取得を防止する必須の技術である．

CAPTCHA は，広く普及した技術だが，多くの研究者によりアクセシビリティの問題が指摘されている．今日主流の CAPTHCA は，歪んだ文字列画像の解釈を利用するが，視覚障害者が解くことは困難である．代替となる方式に，不明瞭な音声を解釈する方式があるが，これは人間でさえ解けないほどに難しい [1], [4], [17]．近年では，視覚障害者らが主体となった CAPTCHA 改善の請願 [12] が行われており，本問題の解決に対する社会的要請の高まりがうかがえる．

本研究の目的は，知覚に関してバリアフリーな

1 明治大学大学院先端数理科学研究科現象数理学専攻
Meiji University, Tokyo 164–8525, Japan
2 筑波技術大学大学院技術科学研究科保健科学専攻
Tsukuba University of Technology, Ibaraki 305–8521, Japan
a) yama3san@meiji.ac.jp
b) ken@cs.k.tsukuba-tech.ac.jp
c) kikn@meiji.ac.jp

CAPTCHA を構成することである．特定の知覚のみの使用に限定されないバリアフリー性と，人間には容易に解けるが，現状のロボットには解答が難しい問題を生成する識別性を必要とする．

1.2 従来方式の問題点

鴨志田らは，ワードサラダと呼ばれるマルコフ連鎖による合成文の不自然さを用いた CAPTCHA を提案した [26]．本稿では，以後これを *KK* 方式と称す．*KK* 方式は，認証の際に，人間の生成した自然文とワードサラダを提示し，利用者にそれらを識別させる．*KK* 方式では，テストをテキストのみで提示可能であり，利用者に要求される知識もその使用言語と一般常識のみである．このように *KK* 方式は，バリアフリー性を満たす興味深い試みの 1 つである．

しかしながら，本稿では，*KK* 方式の問題新規性と識別性に関して，2 つの問題点を指摘する．

(1) 自然文の収集困難性：鴨志田らが論文中で述べているように，ワードサラダは優れた多様性を持つので，自動作問に適している．しかし，人間の記述した多くの文章を必要とする．多様な自然言語データベース（コーパス）を準備しなければならない．

小さなコーパスを用いた場合は，テストとして提示される自然文が使いまわされてしまう．これは，テストに用いた文章を収集し，使いまわされた文を自然文として解答する攻撃に脆弱である．

(2) 検索エンジンを用いた攻撃への脆弱性：新規な問題文を生成するためには，分量の豊富な公開文章をコーパスにする方法が考えられる．しかしこの場合は，検索エンジンを用いた攻撃に対する安全性が問題となる．

自然文は，コーパスとして用いた文章の切り抜きであるため，一般の検索エンジンの検索結果に基づいて，容易に識別できる．

(1)，(2) の問題点は，人間の生成した文を自然文として用いることに起因する．すなわち，これらは 2 章で示す文献 [10], [21] を含む，自然文をテストの一部として提示する方式に共通した問題である．

1.3 本研究の着想

本研究では，これらの問題点を解決するため，自然文がテストに含まれない方式の構築を狙う．具体的には，*KK* 方式の自然文に相当する文にも，マルコフ連鎖により合成したワードサラダを用いる．この方式は，自然文の利用に比べて，以下の点で優れている．

- 同一コーパスに対して，多様な文を生成できる．
- 検索エンジンによるコーパスの特定が困難である．

マルコフ連鎖は，そのパラメータである階数を変えることで，異なる性質を持つワードサラダを生成する．提案方式では，階数の異なるワードサラダの間に存在する「文としての違和感や自然さ」の差を人間に識別させる．階数によって多くの文を合成できるので，(1) の自然文の問題点は存在しない．(2) の検索エンジン攻撃に対しては，ワードサラダのみを使用するので，検索エンジンのデータベースと完全一致する確率は低く，問題点の解決が期待できる．

一方で，提案方式は *KK* 方式の自然文とワードサラダの識別に比べ，人間には解きにくいと推測される．そのため，単一の文ごとに自然文やワードサラダのどちらかを評価させるのではなく，階数の異なる 2 つのワードサラダを 1 組にし，そのどちらが相対的に自然（もしくは不自然）かを問う解答方式を採用する．

1.4 本研究の貢献

本稿では，*KK* 方式の分析と提案方式の実装や実験を通して，次のことを明らかにする．

(1) 自然文を狙う攻撃に対する *KK* 方式の脆弱性　以下に示す攻撃が，鴨志田らの検討した攻撃方式に比べて強力であることを示す．これらの攻撃により，*KK* 方式の安全性が危殆化することを示す．

- 自然文とワードサラダの多様性の差を用いた攻撃：過去に出題された問題文と同一のものが提示された場合，それを自然文として解答する攻撃．
- 検索エンジンを用いた攻撃：自然文とワードサラダを検索し，コーパスを検出した方を自然文として解答する攻撃．

(2) 自然文として扱う合成文の生成に用いる最適な階数　提案方式の自然文に相当するワードサラダを，様々な階数のマルコフ連鎖から合成し，それぞれについて，生成文の多様性や検索エンジンを用いた攻撃に対する安全性を検討する．また，人間による異なる階数で合成したワードサラダの識別能力を，実験により調査する．これらの結果をもとに，安全性と人間による解答容易性をあわせ持つ，最適な階数を決定する．

(3) *KK* 方式に対する提案方式の優位性　提案方式と *KK* 方式の比較を，*F*-値（3.5 節）で定量化する．

本稿の構成は，次に示すとおりである．2 章では，関連研究の紹介を行う．3 章では，準備として，本稿で使用する用語や要素技術，さらに *KK* 方式の概要を示す．4 章において，その問題を解決する方式を提案する．5 章では，提案方式と *KK* 方式を実験的に評価する．6 章では，実験結果から考察を行う．また，提案方式と *KK* 方式の比較を行う．7 章では，本研究を結論づける．

2. 関連研究

ロボットの性能向上にともない，既存の画像 [2], [5]/音声 [3], [18]/クイズ [11] 型 CAPTCHA に対する多くの攻撃成功例が報告されている．このため近年では，ロボットに解答困難と考えられている様々な AI（Artificial In-

telligence）問題を利用したCAPTCHAが提案されている [8], [14], [15], [22].

文意や文脈を解釈する問題（文意文脈解釈問題と称する）は，バリアフリー性を持つAI問題の代表例である．フィッシングメールの識別という限定的な環境でさえ，人間にのみ識別可能な文章の存在がParkら [13] により指摘されている．

文意文脈解釈問題を利用したCAPTCHAは，数多く研究されている．Yamamotoら [21] は，自然文と機械翻訳を繰り返し適用した文との間で人間が感じる違和感をCAPTCHAに用いた．同様に，鴨志田ら [26] は，自然文とマルコフ連鎖による合成文の識別を利用している．Christopher [10] は，複数の文の中から内容が関連するものとしないものを選択させる方式を提案した．またGotoら [6] は，垣根効果を利用した文字削除と，似た音への文字置換を施した文を提示し，元の文を解答させる方式を提案した．

文意文脈解釈問題以外にも，バリアフリー化の試みはされている．Holmanら [7] は，身近な事物を画像と音声の両方で提示し，そのいずれによっても解答可能とする方式を提案した．Shirali-Shahrezaら [16] は，利用者に提示した事物を視覚/聴覚で認識させ，発話により解答させる方式を提案した．

我々の研究グループでは，文意文脈解釈問題を用いたCAPTCHAを検討している．既存のバリアフリーな方式との違いは，CAPTCHAに要求される自動作問性と問題の新規性を満たす点である．既存研究では，この点についての詳細な検討や評価が行われていない．特に，文意文脈解釈問題を用いる既存研究については，必ず自然文を必要とするため，1.2節で指摘した*KK*方式と同様の問題が存在する．

以下に，我々の研究グループによる関連研究を示す．子音交替を用いた方式 [20], [24] では，公開文章を用いることで問題の新規性を満たしつつ，検索エンジンを用いた攻撃にある程度耐性のあるCAPTCHAを提案した．この方式の定性的な評価結果として，子音交替により仮名に開かれる問題文が分かりにくいとの意見があった．このため，子音交替を使用しない方式も検討している．文献 [25] では，ワードサラダを自然文として用いるアイディアの可能性を探るため，異なる階数で合成したワードサラダを被験者に1文ずつ提示し，その自然さについて主観的な評価を調査した．文献 [11] [*1]では，自然文を使用しないワードサラダ識別型CAPTCHAの一案を示した．本稿は，文献 [11] に対して，より人間に解きやすい解答方式への改善と，安全性と人間の正答率についての実験と分析を加えている．

[*1] この論文は，本稿の初版となるIEEE SMC 2014のものである．

3. 準備

3.1 マルコフ連鎖モデル

N をマルコフ連鎖の階数とする．N 階マルコフ連鎖は，直前 N 個の状態に依存して次の状態が決定される確率過程である．状態を表す確率変数 $X_0, \ldots, X_n, X_{n+1}$ について，X_{n+1} の生起確率は，その直前の N 個のみの条件付き確率で，

$$P(X_{n+1} = x | X_n = x_n, \ldots, X_0 = x_0) = P(X_{n+1} = x | X_n = x_n, \ldots, X_{n-N+1} = x_{n-N+1})$$

と与えられる．

N 階マルコフ連鎖による文章合成は，コーパスに形態素解析 [9] を適用して抽出した形態素 N-gramと，そこから連鎖する $N+1$ 番目の形態素の頻度情報からなるマルコフ連鎖モデルを構築して行う．本稿では，階数に幅を持たせたマルコフ連鎖モデルも対象とし，$N_L \leq N \leq N_H$ となる整数 N を階数とする場合は $[N_L, N_H]$ と表す．

3.2 *Hum* と *Spam*

マルコフ連鎖モデルにより合成された文を，ワードサラダと呼ぶ．特に文章合成に使用した階数 N を強調する際は，N 階ワードサラダと称す．

本稿では，*Hum* を「意味論的自然文」，*Spam* を「意味論的不自然文」と定義する．

提案方式では，*Hum* と *Spam* の双方にワードサラダを用いる．*Hum* を合成するマルコフ連鎖モデルの階数 N_{Hum} と *Spam* の階数 N_{Spam} は，$N_{Hum} > N_{Spam}$ の関係を満たす．

提案方式における *Hum* と *Spam* の「自然さ」の尺度は，1つの問題を構成する *Hum* と *Spam* の組における相対的なものであることに注意を要する．

3.3 生成文の多様性とコーパスの多様性

与えられたコーパスから N 階マルコフ連鎖モデルを構築し，それによって生成された W_A 個のワードサラダのうち，重複を除いて互いに異なるものが W_U 個あるとする．本稿では，$100 \times W_U/W_A[\%]$ を生成文の多様性と定義し，問題新規性の評価指標として扱う．たとえば，生成した100個のワードサラダ中に10個の重複があった場合，その多様性は90%になる．

コーパスに含まれる形態素 M-gram の集合を $\mathcal{D}_M$ とし，形態素1-gramからなる集合を $\mathcal{D}_1$ とする．ある形態素 M-gram $A(= a_n a_{n-1} \ldots a_{n-M+1}) \in \mathcal{D}_M$ から連鎖する $M+1$ 番目の形態素の候補集合 $\mathcal{C}_{(M,A)}$ は，

$$\mathcal{C}_{(M,A)} = \{c \in \mathcal{D}_1 \mid P(X_{n+1} = c | X_n = a_n, \ldots, X_{n-M+1} = a_{n-M+1}) > 0\}$$

となる．すなわち，生起確率が0より大きな候補の総数を

表す．本稿では $C_M = \sum_{A \in \mathcal{D}_M} |\mathcal{C}_{(M,A)}|/|\mathcal{D}_M|$ を，形態素 M-gram でのコーパスの多様性と定義する．

3.4 *KK* 方式

KK 方式のアルゴリズムを以下に示す．

KK 方式のアルゴリズム [26]

(1) コーパスから N 階マルコフ連鎖モデルを作る．
(2) 自然文を h 個，ワードサラダを s 個，計 z 個の文をランダムな順で利用者に与える．
(3) 利用者は z 個の文を，それぞれ *Hum* か *Spam* に分類して解答する．
(4) 正答数 k を求め，$k \geq$ 閾値 θ ならば利用者を受理，そうでなければ拒否する．

KK 方式の特徴は，問題文に自然文を使用している点である．このため *KK* 方式では，検索攻撃を避けるため，秘匿文章をコーパスにすることを推奨している．

3.5 安全性定義と評価方法

本稿では，鴨志田ら [26] の安全性定義を利用する．

評価指標（*FRR*，*FAR*，*F*-値）

X を出題文を表す確率変数，Y を解答を表す確率変数，H を *Hum*，S を *Spam* とする．作問者が正答が *Hum* となる文を出題して，利用者が *Hum* と解答する条件付き確率は，$P(Y = H|X = H)$ と表せる．*Hum* と *Spam* を出題する確率はそれぞれ，

$$P(X = H) = \frac{h}{z}$$
$$P(X = S) = \frac{s}{z} = 1 - \frac{h}{z}$$

となる．CAPTCHA の成功率は，これらの同時確率で，

$$P(Y = H, X = H) = P(Y = H|X = H)P(X = H)$$
$$P(Y = S, X = H) = P(Y = S|X = H)P(X = H)$$
$$P(Y = H, X = S) = P(Y = H|X = S)P(X = S)$$
$$P(Y = S, X = S) = P(Y = S|X = S)P(X = S)$$

と与える．人間による CAPTCHA 1 問あたりの失敗率を，

$$P_q = P(Y = S, Y = H) + P(Y = H, X = S)$$

とする．

ここで，CAPTCHA z 問により構成された認証方式を考える．人間による CAPTCHA の正答数が $k < \theta$ となる確率を，人間拒否率 FRR（False human Rejection Rate）と定める．また，ロボットによる CAPTCHA 1 問あたりの成功率を P_m とし，その正答数が $k \geq \theta$ となる確率を，機械受入率 FAR（False machine Acceptance Rate）と定める．すなわち，FRR および FAR は，1 回あたりの確率 P_q および P_m となる事象が z 回中 k 回発生し，かつ閾値 θ と k が前述の関係を満たす確率である．よって FRR および FAR を，二項分布で

$$FRR = \sum_{k=\theta}^{z} \binom{z}{k} P_q^k (1 - P_q)^{z-k} \tag{1}$$
$$FAR = \sum_{k=\theta}^{z} \binom{z}{k} P_m^k (1 - P_m)^{z-k}$$

と与える．

本稿では，*KK* 方式と提案方式の比較を容易にするため，$\theta = 1$，$z = 1$ での FRR と FAR から，次の F-値

$$F = \frac{2 \cdot (1 - FAR) \cdot (1 - FRR)}{(1 - FAR) + (1 - FRR)} \tag{2}$$

を用いる．

ツールを用いた攻撃に対する安全性

ツールによる特定の処理を表す確率変数を W とする．処理の具体例としては，MS-WORD による文章校正 [26] や検索エンジンによるコーパスの検出がある．処理が行われる事象を $W = t$ とすれば，その確率 $P(W = t)$ は，

$$\begin{aligned} P(W = t) &= P(W = t, X = S) + P(W = t, X = H) \\ &= P(W = t|X = S)P(X = S) \\ &\quad + P(W = t|X = H)P(X = H) \end{aligned} \tag{3}$$

となる．このとき，入力が *Spam* である確率は，ベイズの定理から，

$$P(X = S|W = t) = \frac{P(W = t|X = S)P(X = S)}{P(W = t)} \tag{4}$$

となる．同様に，処理が行われない事象を $W = f$ とすれば，その確率 $P(W = f)$ は，

$$\begin{aligned} P(W = f) &= P(W = f, X = S) + P(W = f, X = H) \\ &= P(W = f|X = S)P(X = S) \\ &\quad + P(W = f|X = H)P(X = H) \end{aligned} \tag{5}$$

となる．このとき，入力が *Hum* である確率は，

$$P(X = H|W = f) = \frac{P(W = f|X = H)P(X = H)}{P(W = f)} \tag{6}$$

となる．したがって，ロボットによる解答 Y_w を，$W = t$ のとき，

$$Y_w = \begin{cases} S \text{ w./p. } P(X = S|W = t) \\ H \text{ w./p. } P(X = H|W = t) \end{cases} \tag{7}$$

$W = f$ のとき，

$$Y_w = \begin{cases} S \text{ w./p. } P(X = S|W = f) \\ H \text{ w./p. } P(X = H|W = f) \end{cases} \tag{8}$$

と定めることで，FAR を最大化できる．したがって，CAHTCHA 1 問あたりの正答率は，

$$P_{mw} = P(Y_w = S, X = S) + P(Y_w = H, X = H) \tag{9}$$

となる．ただし，

$$P(Y_w = S, X = S) = P(Y_w = S|W = t)P(W = t|X = S) + P(Y_w = S|W = f)P(W = f|X = S) \quad (10)$$

$$P(Y_w = H, X = H) = P(Y_w = H|W = t)P(W = t|X = H) + P(Y_w = H|W = f)P(W = f|X = H) \quad (11)$$

である．

4. 提案方式

4.1 提案方式の概要

KK 方式と提案方式の違いは，(1) *Hum* の生成方法と (2) 解答方式である．

(1) *Hum* の生成方法：提案方式では，*Hum*，*Spam* ともにマルコフ連鎖で合成する．ただし，それぞれに使用する階数は $N_{Hum} > N_{Spam}$ を満たす．

Hum にワードサラダを用いることは，以下の点で自然文の利用に比べて優れている．よって，ロボットに対する *FAR* の改善が期待できる．

- 同一コーパスに対して，より高い生成文の多様性を持つ．
- 検索エンジンによるコーパスの特定が困難である．

提案方式は，上記の利点により，公開文章をコーパスとして使用できる．事前にある程度大きなコーパスを準備しておけば，問題を生成しながらコーパスを公開文章から収集し，マルコフ連鎖モデルを更新できる．定期的なモデルの更新により，生成文の多様性を維持することができる．

(2) 解答方式：ワードサラダは，自然文に比べて違和感の強い文章が生成されやすいため，人間に対する *FRR* の悪化が懸念される．その対策として，提案方式では 1 問ごとに *Hum* と *Spam* を両方提示し，そのどちらが相対的に自然（もしくは不自然）かを問う方式を取り入れる．

図 1 に提案方式の作問例を示す．作問例 1 は，*Hum* 単体では自然さを感じることはできない．作問例 2 でも，人によっては違和感を感じるかもしれない．しかし，*Spam* との相対比較にすることで，解答が容易になっている．

4.2 方式定義

提案方式のアルゴリズムを以下に示す．

提案方式のアルゴリズム

(1) $N_{Hum} > N_{Spam}$ となる 2 つの階数を選択し，コーパスからそれぞれのマルコフ連鎖モデルを作る．
(2) *Hum* と *Spam* を z 組作成する．
(3) 各組に対して，*Hum* と *Spam* をランダムに選択肢 A と B に割り当て，利用者に提示する．利用者に，各組についてより自然（*Hum*），もしくは不自然（*Spam*）な選択肢を解答させる．
(4) 正答数 k を求め，$k \geq$ 閾値 θ ならば利用者を受理，そうでなければ拒否する．

問題: より不自然な文を選択してください．

No. 1 ($N_{Hum} = 2, N_{Spam} = 1$)
A 離れとは言い出せなかったが，その彼とほぼ同時に停留所に着くだろ
B 菜ならんで水を終ると，彼の自転車の間，つらくなって，一声であっ

No. 2 ($N_{Hum} = 3, N_{Spam} = 1$)
A 三歩下がり，邦子は気にいっていた．片方の壁には食器棚があり，
B 東亜戦争が，そして転んでいる海岸にあおむけになり，御隠居と戦っ

No. 3 ($N_{Hum} = 4, N_{Spam} = 1$)
A 亜紀子はきめた．水の表面がきらきらと輝きながら，小さく揺れて
B 加えた．入口までも四歳にあたえた．車を覚えにいるとすごいの気持

解答: 全て B が *Spam*.

図 1 提案方式による作問例
Fig. 1 Sentences synthesized by our proposal.

コーパスは，問題新規性を満たすため，一定の作問数ごとに更新が必要になる．コーパスのサイズに依存して，十分な問題新規性が維持される作問数は変化するため，事前に多様性を実験的に確認することが望ましい．

5. 評価

5.1 評価項目

提案方式の有効性を検証するため，次の実験を行う．

実験 1 生成文の多様性の評価
実験 2 検索エンジンを用いた攻撃による *Hum* と *Spam* の識別性の評価
実験 3 人間による評価

5.2 実験方法

5.2.1 共通設定と表記方法

次に，実験に共通する設定を示す．

- 青空文庫 [23] に登録されている 5 種類の現代語仮名遣いの文章を，まとめて 1 つのコーパスとして扱う．**表 1** にその特徴を示す．
- 提案方式で *Spam* に用いるワードサラダの階数を $N_{Spam} = 1$ とする．*Hum* に用いるワードサラダの階数 N_{Hum} は，固定階数である 1, 2, 3, 4, 5 と，幅を持つ階数 $[N_L, N_H] = [1, 2], [1, 3], [2, 3], [2, 4], [3, 4], [3, 5], [4, 5]$ を用いて評価する．

表 1 実験に用いたコーパスの特徴 (文字数, 行数) = (80783, 5248)
Table 1 Features of our corpus; (Number of characters, Lines) = (80783, 5248).

N-gram	1	2	3	4	5	6	7
Number of Unique Words	7,893	34,469	60,790	73,632	77,532	77,526	76,395
Diversity of the Corpus (C_N)	4.403	1.785	1.231	1.075	1.023	1.008	1.002

- *KK* 方式を，$N_{Hum} = 7$ と $N_{Spam} = 1$ のワードサラダ識別問題として扱う．表 1 から，7-gram におけるコーパスの多様性は $C_{N=7} \approx 1$ なので，本稿では 7 階ワードサラダを自然文として扱う．
- 生成するワードサラダの文字数は，30 から 40 文字の範囲とする．

グラフの表記について，その注意点を示す．階数をパラメータとして扱う場合，N と $N_{diff} = N_H - N_L$ を使用する．固定階数の場合，$N_{diff} = 0$ とする．幅を持つ階数の場合，$N = N_H$，$N_{diff} = N_H - N_L$ とする．

5.2.2 実験 1（多様性）

マルコフ連鎖モデルからワードサラダを，階数ごとに 50,000 個生成し，生成文の多様性を確認する．

5.2.3 実験 2（検索エンジンによる識別性）

次の手順を，$N_{Hum} = 2, 3, 4, 5$，$N_{diff} = 0, 1, 2$ について，それぞれ 10 回行う．

(1) N_{Hum} 階マルコフ連鎖モデルから，ワードサラダ 10 個を *Hum* として合成する．1 階マルコフ連鎖モデルから，ワードサラダ 10 個を *Spam* として合成する．
(2) *Hum* と *Spam* を順番に 1 つずつ取り出して組を構成する．すべての組に対して，次の処理を行う．
   - Yahoo!検索エンジンにそれぞれを問い合わせる．
   - 検索結果の上位 10 件にコーパスが含まれている場合，検索エンジンがコーパスを検出したと判断する．コーパスの検出が成功した場合，さらに次の判定を行う．
   - 次のいずれかの条件を満たした場合，検索エンジンにより *Hum* と *Spam* を正しく識別できたとする．
     – *Hum* のみコーパスが検出された．
     – *Hum* の方が *Spam* より上位検索結果で，コーパスの検出がされた．

   検索では各文について通常方式と完全一致方式の両方を実行し，上位となる方を利用する．

利用する検索エンジンは，Yahoo!，Google，Bing を候補にあげ，事前調査し Yahoo!を選択した．Bing 検索は，ワードサラダからコーパスを検出する確率が低いため除外した．Google 検索は，プログラムによる連続した検索ができなかったため，効率の観点から除外した．

5.2.4 実験 3（人間による評価）

次の手順で生成したエクセルファイルを用いて，テストを実施した．被験者は，男性 13 名，女性 3 名の日本人 16 名である．年齢構成は 18 歳から 65 歳であり，視力の状態は全盲（光覚・手動弁・指数弁）が 2 件，弱視が 3 件，晴眼が 11 件であった．

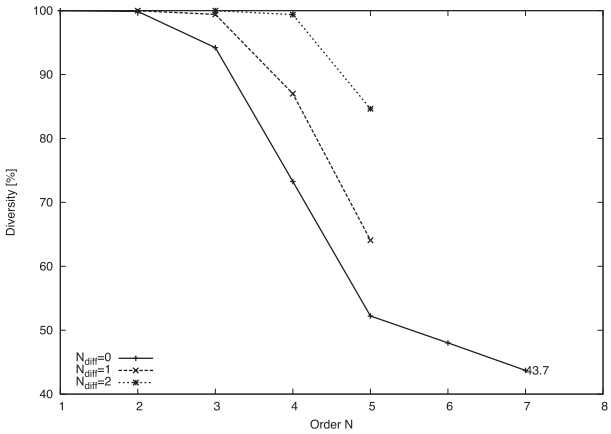

図 2 N 階ワードサラダの多様性
Fig. 2 Diversity of sentences generated by Markov chain.

(1) 実験 2 で生成した *Hum* と *Spam* の組を，各 *Hum* の階数ごとに，無作為に 10 個選択する．
(2) 選択肢として A と B の 2 つを用意し，各組ごとに *Hum* と *Spam* を無作為に割り当てる．

5.3 実験結果

5.3.1 実験 1（多様性）

図 2 に，N 階ワードサラダの多様性を示す．

固定階数（$N_{diff} = 0$）のグラフから，自然文（$N = 7$）に比べて，ワードサラダの多様性が十分高いことが分かる．特に $N < 4$ で顕著である．$N = 4$ を境に傾きが急峻になるが，以降は N の増加に対して緩やかに遷移している．この理由は，表 1 のコーパスでは $C_{N \geq 4} \approx 1$ のためと推測される．

幅を持つ階数（$N_{diff} = 1, 2$）で生成したワードサラダの多様性は，固定階数 N_H と N_L で生成したものの値に対して，N_L よりの中間位置にあることから，N_L の影響が強く出たと推測される．

5.3.2 実験 2（検索エンジンによる識別性）

表 2 に，検索エンジンによるワードサラダ単文ごとのコーパス検出率を示す．また，**図 3** に，検索エンジンによる *Hum* と *Spam* の識別結果を示す．図中の縦線は，$\pm 1\sigma$ の幅を示す．

固定階数（$N_{diff} = 0$）のグラフから，ワードサラダどうしの識別は，自然文（$N = 7$）とワードサラダの識別に比べて，検索エンジンには困難である．検索エンジンを用い

表 2 検索エンジンによるコーパス検出確率
Table 2 Conditional probabilities of sentence to be detected.

Order N	1	2	3	4	5	7	[1,2]	[1,3]	[2,3]	[2,4]	[3,4]	[3,5]	[4,5]
$P(W=t\|X=x)^{\dagger}$	12	19	44	78	85	89	6	9	25	33	56	58	75

†: If $N=1$, then $x=S$. Otherwise, $x=H$.

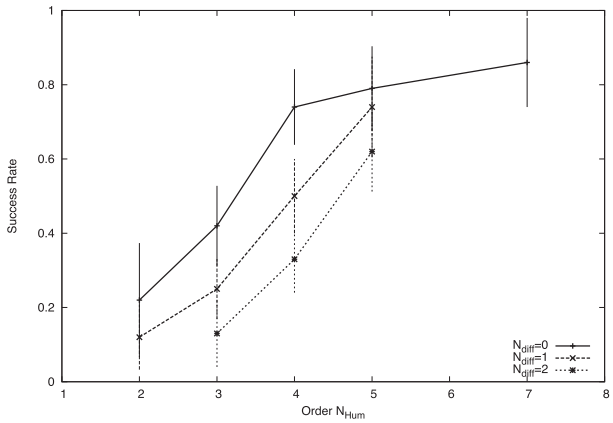

図 3 Yahoo! 検索エンジンによる *Hum* と *Spam* の識別能力
Fig. 3 Distinguishability rate by Yahoo! search engine.

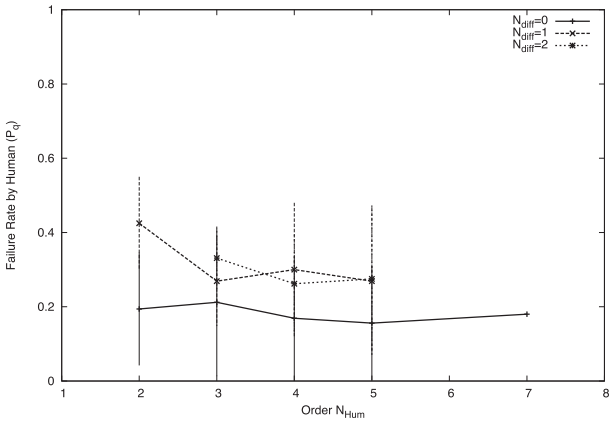

図 4 人間による *Hum* と *Spam* の識別結果（失敗率 P_q）
Fig. 4 Distinguishability rate by human.

た攻撃者は，*KK* 方式に相当する $N_{Hum}=7$, $N_{Spam}=1$ の 2 つを 86%で識別可能だが，$N_{Hum}=2$, $N_{Spam}=1$ になると 22%しか識別できない．

N_{Hum} に幅を持つ階数（$N_{diff}=1,2$）と固定階数（N_H, N_L）を用いた場合を比較する．検索エンジンによる $N_{Hum}=[N_L,N_H]$ と $N_{Spam}=1$ のワードサラダとの識別結果は，N_L よりの中間位置にある．実験 1 の結果と同様に，N_L の影響が強く出たと推測される．

5.3.3 実験 3（人間による評価）

図 4[*2]に，主観実験により取得した，人間の CAPTCHA 1 問あたりの失敗率 P_q を示す．図中の縦線は，$\pm 1\sigma$ の幅を示す．

固定階数（$N_{diff}=0$）のグラフから，*KK* 方式（$N_{Hum}=7$）の結果 18%に対して，提案方式の結果は 15.6–21.2%であり，階数に依存せずほぼ一定の P_q となることが分かった．また，幅を持つ階数（$N_{diff}=1,2$）では固定階数に比べて若干高い P_q が出ているが，$N_L=1$ の場合を除き，階数の変化に対してはあまり影響が見られなかった．

著者らは，人間による識別結果も実験 1 や 2 と同様に，階数変動の影響を強く受けると予想していた．実験 3 の結果は，人間による文の自然さの認識力が，想像以上であることを示している．

[*2] $N_{Hum}=7$ のデータは，鴨志田らの結果から導出したものであるが，平均値のみをプロットしている．また解答方式は，単文ごとに *Hum* と *Spam* を識別するものである．

6. 考察

6.1 *KK* 方式の脆弱性

KK 方式について，鴨志田らの検討したランダム推測攻撃とワード攻撃 [26] に加えて，本稿で示した，生成文の多様性や検索エンジンを用いた攻撃についての安全性を検討する．

ランダム推測攻撃

h,s の値を知る攻撃者によるランダム推測の攻撃成功率 P_{mr} は，次のように示される．

$$\begin{aligned} P_{mr} &= P(Y=S,X=S)+P(Y=H,X=H) \\ &= P(Y=S)P(X=S)+P(Y=H)P(X=H) \\ &= \left(\frac{s}{z}\right)^2+\left(\frac{h}{z}\right)^2 \end{aligned}$$

他の攻撃については，$s=5$，$h=15$ を例にあげて計算方法を示す．

ワード攻撃

ワード攻撃成功率 P_{mw} の計算方法を示す．問題文 X が MS-WORD 2007 による文章校正を受ける事象を $W=t_w$，校正を受けない事象を $W=f_w$ とすれば，文献 [26] より $P(W=t_w|X=S)=0.24$, $P(W=t_w|X=H)=0$ となる．$P(X=S)=0.25$, $P(X=H)=0.75$ なので，式 (3) より，$P(W=t_w)=0.06$, $P(W=f_w)=0.94$ となる．この分類器では，文章校正がされた場合には必ず *Spam* と解答する．そうでなければ式 (6) より，分類器は $P(X=H|W=f_w)=0.798$ の確率で *Hum*, $P(X=S|W=f_w)=0.202$ の確率で *Spam* と解答する．式 (7)，(8) の議論より，式 (10)，(11) において $P(Y_w=H,X=H)=0.798$，$P(Y_w=S,X=S)=0.394$ となる．よって，式 (9) から $P_{mw}=0.697$ となる．

生成文の多様性の差を用いた攻撃

この攻撃では，過去に出題された問題文を収集し，新た

に提示された問題が過去に出題されたものと一致するかどうかを調べ，その結果を攻撃に利用する．*Hum* と *Spam* の多様性が異なる場合に，有効な攻撃手段である．

生成文の多様性の差を用いた攻撃の成功率 P_{md} の計算方法を示す．問題文 X が過去に出題されたものと一致する事象を $W = t_d$，一致しない事象を $W = f_d$ とすれば，実験 1 の $N = 1,7$ の結果より $P(W = t_d|X = S) = 0$，$P(W = t_d|X = H) = 0.563$ となる．この分類器では，$W = t_d$ であれば必ず *Hum* と解答する．そうでなければ式 (6) より，$P(X = H|W = f_d) = 0.567$ の確率で *Hum* を，$P(X = S|W = f_d) = 0.433$ の確率で *Spam* と解答する．以降はワード攻撃と同じ議論により，$P_{md} = 0.716$ が計算される．

検索エンジンを用いた攻撃

この攻撃は，*Hum* と *Spam* の間に存在する検索結果の違いを利用する．

検索エンジンを用いた攻撃の成功率 P_{ms} の計算方法を示す．問題文 X が検索エンジンによりコーパスを特定された事象を $W = t_s$，特定されない事象を $W = f_s$ とする．表 2 から $N = 1$ のデータを *Spam*，$N = 7$ のデータを *Hum* とすれば，$P(W = t_s|X = S) = 0.12$，$P(W = t_s|X = H) = 0.89$ となる．この分類器では，$W = t_s$ であれば式 (4) より $P(X = H|W = t_s) = 0.957$ の確率で *Hum* を，$P(X = S|W = t_s) = 0.043$ の確率で *Spam* と解答する．そうでなければ，式 (6) より $P(X = H|W = f_s) = 0.273$ の確率で *Hum* を，$P(X = S|W = f_s) = 0.727$ の確率で *Spam* と解答する．以降はワード攻撃と同じ議論により，$P_{ms} = 0.823$ が計算される．

攻撃方式の比較

図 5 に，$P(X = S) = s/z$ について，*KK* 方式に対する各攻撃方式の成功率を示す．本稿で検討した 2 つの攻撃手法は，鴨志田らの検討したワード攻撃より，高い性能を示している．

攻撃手法により，*Hum* と *Spam* のどちらの検出を得意とするかは異なる．そのため s，h の比率により攻撃成功確率は異なるが，それぞれの最低値は $(P_{mr}, P_{mw}, P_{md}, P_{ms}) = (0.500, 0.567, 0.683, 0.796)$ となることから，*KK* 方式は検索エンジンを用いた攻撃に最も脆弱である．

P_{md} については，使用するコーパスとそこから生成する文の数により，強い影響を受けることに注意を要する．たとえば，同一コーパスからの作問数が増えたり，小さいコーパスを利用した場合は，P_{md} の値は上昇してしまう．

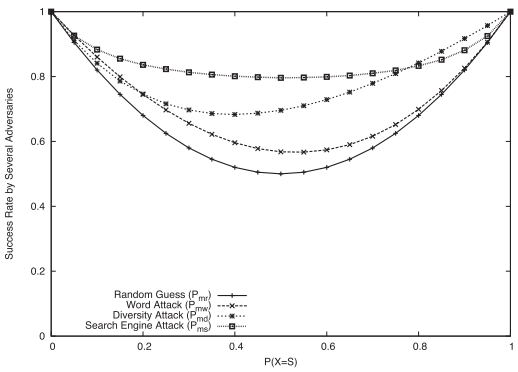

図 5 *KK* 方式に対する $P(X = S)$ ごとの攻撃成功率

Fig. 5 Attack success rate given known probability of *Spam* $P(X = S)$ in *KK* scheme.

6.2 最適な N_{Hum} の決定と既存方式との比較

実験 1，2 の結果と 6.1 節の検討より，検索エンジンを用いた攻撃が最も高い成功率を示すことから，この結果を提案方式と *KK* 方式の FRR として用いる．**図 6** に，表 2 と 6.1 節で示した計算法から，CAPTCHA 1 問あたりの検索エンジンを用いた攻撃の成功率を示す．このグラフでは，図 4 の結果より，最適な N_{Hum} の候補となる固定階数のデータのみを示している．なお，同様の理由により，以後の検討は固定階数の場合のみで行う．

図 4 と図 6 から得た FRR と FAR より，総合的な指標として，式 (2) に示される F-値を用いて，提案方式と *KK* 方式を比較する．他の既存方式については，論文ごとに提示された FRR と FAR から F-値を計算する．

表 3 に，既存方式と提案方式の比較結果[*3]を示す．表 3 のバリアフリーな方式とは，特定知覚に依存せず，問題の認識と解答ができるものと定める．テキストを提示する形式の CAPTCHA は，聴覚障害者は視認により，視覚障害者は彼らの使用する一般的な補助ソフトであるスクリーンリーダにより，問題への対処ができる．商用方式については，バリアフリー性を満たさないため，参考データとしての扱いに留める．

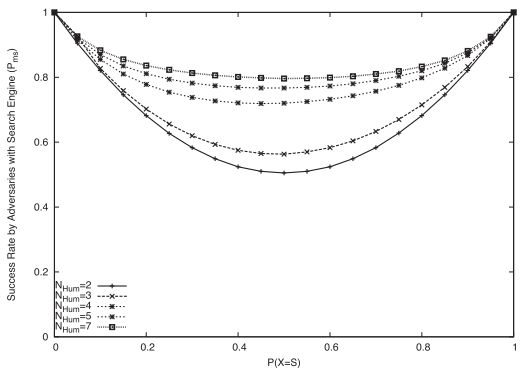

図 6 提案方式に対する $P(X = S)$ ごとの検索エンジンを用いた攻撃成功率

Fig. 6 Attack success rate given known probability of *Spam* $P(X = S)$ in our proposal.

[*3] 商用方式の FRR は，FAR と比べて調査時期が古いことに注意を要する．特に音声型 CAPTCHA は，その難化が指摘されている [12], [17], [24] ため，これらの FRR が悪化している可能性がある．

表 3 CAPTCHA 1 問あたりにおける既存方式と提案方式の比較
Table 3 Comparison between conventional schemes and our proposal.

Scheme	N_{Hum}	FRR	FAR	F-ratio	Barrier-free?
Visual-eBay†1	—	0.07	0.514	0.64	No†4
Visual-reCapcha†1	—	0.25	0.223	0.76	No†4
Visual-Yahoo†1	—	0.12	0.053	0.91	No†4
Audio-eBay†2	—	0.37	0.829	0.27	No†5
Audio-reCapcha†2	—	0.53	0.015	0.64	No†5
Audio-Yahoo†2	—	0.32	0.455	0.61	No†5
[21]†3	7	0.00	0.796	0.35	Yes
KK [26]	7	0.180	0.796	0.33	Yes
Our Proposal	2	0.194	0.505	0.61	Yes
	3	0.212	0.563	0.56	
	4	0.169	0.720	0.42	
	5	0.156	0.767	0.37	

†1: The values of FRR and FAR are referred from [4] and [2], respectively.
†2: The values of FRR and FAR are referred from [4] and [3], respectively.
†3: The value of FRR is referred from the results of informal experiments described in [21].
†4: For the hearing impaired.
†5: For the visually impaired.

表 3 の結果から，提案方式は $N_{Hum} = 2$ の場合に F-値 0.61 で最適となる．バリアフリーな方式としては，提案方式が最も良い F-値を示した．

6.3 Gap Amplification

CAPTCHA 1 問あたりの FRR と FAR において，$FAR > FRR$ が成立する場合は，Gap Amplification [19] により安全性を強化できる．

提案方式において，6.2 節で導出した最適条件 $H_{Hum} = 2$, $H_{Spam} = 1$ の問題を，20 問出題する場合を考える．式 (1) における (z, P_q, P_m) はそれぞれ (20, 0.194, 0.505) となることから，正答数の閾値 θ をパラメータとした CAPTCHA 20 問あたりの FRR と FAR は，二項分布より**図 7** のように計算できる．$FRR = FAR$ となる ERR（Equal Error Rate）の条件において，提案方式は $\theta = 6, 7$ の付近で，FRR と FAR をともに約 10%に改善できる．

6.4 コーパスの違いによる生成文の多様性

図 8 に，異なるコーパスを用いた場合の生成文の多様性を示す．各コーパスの特徴は，異なり語数については**図 10** に，コーパスの多様性（C_N）については**図 9** に示す．実験では，文章は各々が異なる特徴を持つことを想定し，意図的に特徴の異なるコーパスを用いた．図 8 は，コーパス 0–4 ごとに，N 階ワードサラダを 50,000 個ずつ生成したときの生成文の多様性の分布である．なお，5 章で用いたコーパスは，コーパス 0–4 を 1 つにまとめたものである．

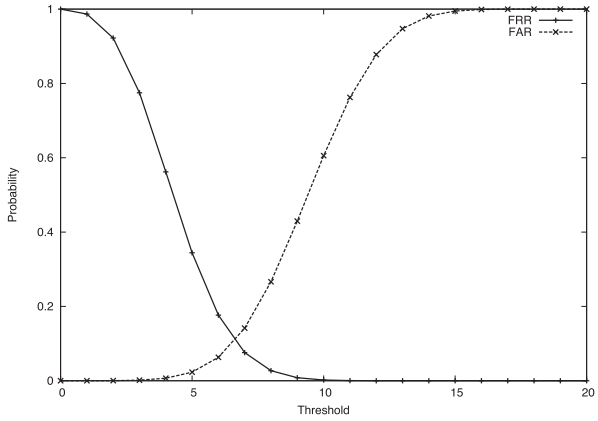

図 7 提案方式（$N_{Hum} = 2$, $N_{Spam} = 1$）における FAR と FRR の分布
Fig. 7 Probability distributions of FAR and FRR of our proposal ($N_{Hum} = 2, N_{Spam} = 1$).

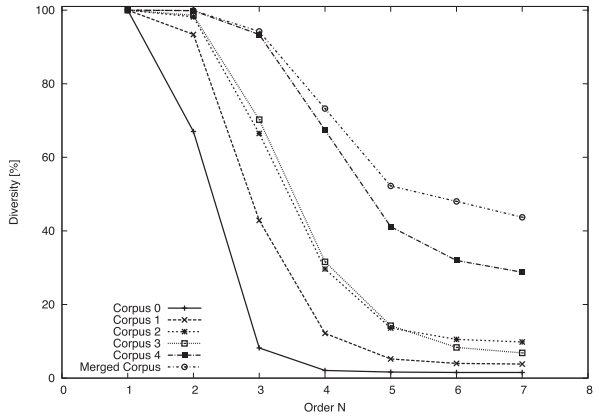

図 8 異なるコーパスから生成された N 階ワードサラダの多様性（$N_{diff} = 0$）
Fig. 8 Diversity of sentences generated by different corpora ($N_{diff} = 0$).

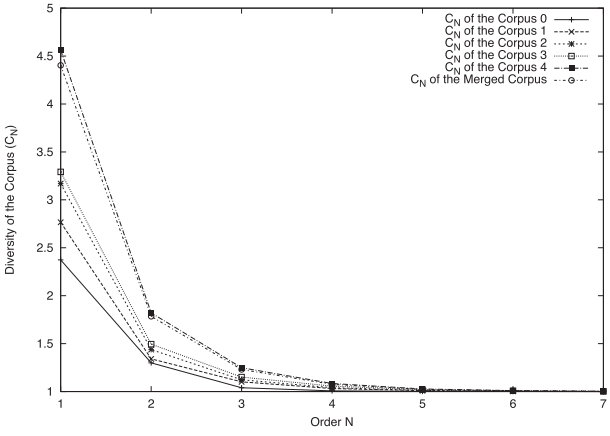

図 9 コーパスの多様性
Fig. 9 Diversity of our corpora.

図 8 から，階数 N に対する生成文の多様性は，コーパスの異なり語数と C_N の影響を受けることが分かった．特に C_N の影響は，N が増加するほど顕著に表れた．したがって，生成文の多様性を確保するためには，分量の大き

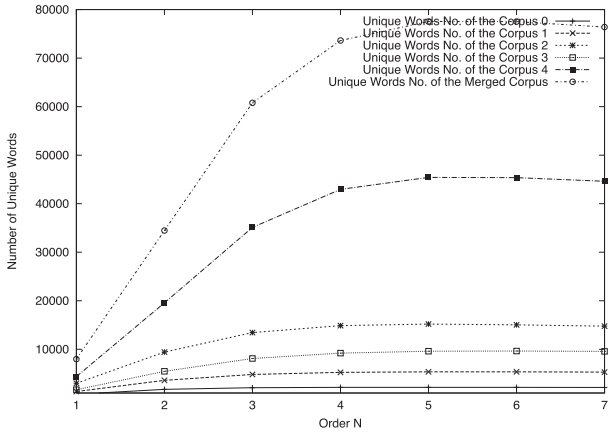

図 10 コーパスの異なり語数
Fig. 10 Number of unique words of our corpora.

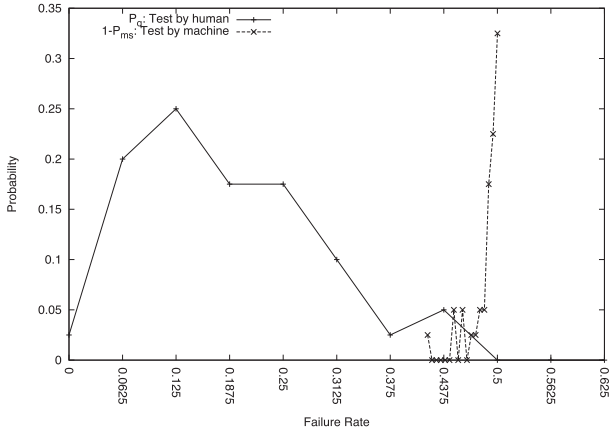

図 11 失敗率（$P_q, 1-P_{ms}$）の度数分布
Fig. 11 Frequency distribution of failure rate ($P_q, 1-P_{ms}$).

いコーパスの採用や，小さい N 階ワードサラダの利用が効果的である．また，異なり語数と C_N を調整したコーパスの選択をすることで，ある程度は生成文の多様性を制御できると推測する．

6.5 失敗率 $P_q, 1-P_m$ の分布

図 11 に，提案方式における CAPTCHA 1 問あたりの人間と機械の失敗率 $P_q, 1-P_{ms}$ の度数分布を示す．

図 4 の結果から，P_q と N_{Hum} への依存性が小さいので，人間の失敗率 P_q については $N_{diff}=0$ となる 40 問を 1 つの群として扱う．ロボットの失敗率には，最も強力な検索攻撃を代表例とし，$N_{Hum}=2$ における $1-P_{ms}$ を用いる．ロボットの失敗率は，コーパスより生成した 640 問と 6.1 節に示した算出法から，16 問ごとの検索結果をもとに計算した 40 個のデータを用いる．

P_q は，(平均値, 標準偏差) $=(0.183, 0.106)$ となる分布であった．$1-P_{ms}$ は，(平均値, 標準偏差) $=(0.485, 0.018)$ となる分布であった．提案方式の P_q と $1-P_{ms}$ の分布は，重なりが少なく，かつその中央値どうしが離れているため，CAPTCHA としての機能が期待できる．

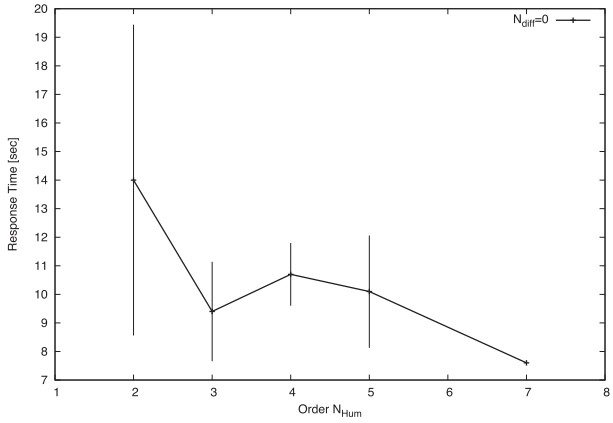

図 12 1 問あたりの利用者の応答時間
Fig. 12 Response time per question.

P_q と $1-P_{ms}$ は，平均値と中央値がほぼ一致した山なりの分布である．したがって，P_q や P_{ms} の平均値を代表値とした複数回試行への二項分布による近似（6.3 節）は，一定の信頼性があると考える．

6.6 課題：利用者の応答時間

図 12[*4]に，実験 3 の正答率が 16 名中 8 番目であった被験者の応答時間を示す．図中の縦線は，$\pm 1\sigma$ の幅を示す．被験者数を増やした信頼性の確保は，今後の課題とする．

ワードサラダ識別型 CAPTCHA は，既存の方式に比べ応答時間が長いことが知られている．提案方式の問題 20 問を用いて CAPTCHA を構成した場合，認証には 200 秒程度の応答時間が見込まれる．一方で，商用 CAPTCHA の平均応答時間は，画像型で 11 秒，音声型で 43 秒程度である [4]．注意を要する点として，FRR が高い方式は，認証を受けるまでに複数の試行が必要になることがある．特に音声型 CAPTCHA は，この理由により，認証までの応答時間が増大しやすい．

KK 方式 [26] と比べて提案方式は，1 問あたり 2 つのワードサラダを読む必要があるので応答時間が長い．特に $N_{Hum}=2$ の場合に顕著であるが，これは実験 3 の順番として $N_{Hum}=2$ を最初に行った影響も考えられる．$N_{Hum}=2$ の後半 5 問に限れば，応答時間の平均は 10.8 秒であり，$N_{Hum}=3,4,5$ の場合と同程度である．

以上の検討から，提案方式の応答時間の改善は，重要な課題であると考えられる．対策としては，ワードサラダの文字数を削減し，利用者の読む文章量を削減する方法がある．

7. おわりに

本稿では，鴨志田ら [26] により提案された，自然文を用いたワードサラダ識別型 CAPTCHA の安全性の問題を指

[*4] $N_{Hum}=7$ のデータは文献 [26] から取得し，平均値のみをプロットしている．

摘し，自然文の収集困難性と検索エンジンを用いた攻撃に対する脆弱性を実験的に示した．

提案方式は，階数の異なるマルコフモデルから生成された2種類のワードサラダ間に存在する「文の自然さ」の差をCAPTCHAに利用する．提案方式では，ワードサラダのみを用いることで，自然文に起因する脆弱性の問題を解決した．また，2種類のワードサラダの比較結果を解答する方式により，人間による正答率の低下を抑制した．本稿では，提案方式と鴨志田らの方式をF-値を指標として比較し，その優位性を示した．

謝辞 本稿作成にあたり，産業技術総合研究所セキュアシステム研究部門の中田亨氏には，終始適切な助言をいただきました．ここに感謝の意を表します．

参考文献

[1] Bigham, J.P. and Cavender, A.C.: Evaluating existing audio CAPTCHAs and an interface optimized for non-visual use, *Proc. SIGCHI Conference on Human Factors in Computing Systems*, pp.1829–1838, ACM (2009).

[2] Bursztein, E., Aigrain, J., Moscicki, A. and Mitchell, J.C.: The End is Nigh: Generic Solving of Text-based CAPTCHAs, *8th USENIX Workshop on Offensive Technologies (WOOT 14)*, San Diego, CA, USENIX Association (online), available from ⟨https://www.usenix.org/conference/woot14/workshop-program/presentation/bursztein⟩ (2014).

[3] Bursztein, E., Beauxis, R., Paskov, H.S., Perito, D., Fabry, C. and Mitchell, J.C.: The Failure of Noise-Based Non-continuous Audio Captchas, *32nd IEEE Symposium on Security and Privacy, S&P 2011*, 22-25 May 2011, Berkeley, California, USA, pp.19–31 (2011).

[4] Bursztein, E., Bethard, S., Fabry, C., Mitchell, J.C. and Jurafsky, D.: How Good Are Humans at Solving CAPTCHAs? A Large Scale Evaluation, *Proc. 2010 IEEE Symposium on Security and Privacy*, pp.399–413, IEEE Computer Society (2010).

[5] Bursztein, E., Martin, M. and Mitchell, J.: Text-based CAPTCHA Strengths and Weaknesses, *Proc. 18th ACM Conference on Computer and Communications Security*, pp.125–138, ACM (2011).

[6] Goto, M., Shirato, T. and Uda, R.: Text-Based CAPTCHA Using Phonemic Restoration Effect and Similar Sounds, *IEEE 38th Annual Computer Software and Applications Conference, COMPSAC Workshops 2014*, pp.270–275 (2014).

[7] Holman, J., Lazar, J., Feng, J.H. and D'Arcy, J.: Developing usable CAPTCHAs for blind users, *Proc. 9th International ACM SIGACCESS Conference on Computers and Accessibility*, pp.245–246, ACM (2007).

[8] Imsamai, M. and Phimoltares, S.: 3D CAPTCHA: A Next Generation of the CAPTCHA, *Proc. Information Science and Applications 2010, ICISA '10*, pp.1–8, IEEE (2010).

[9] Kudo, T.: MeCab: Yet Another Part-of-Speech and Morphological Analyzer, http://mecab.sourceforge.net/ (online), available from ⟨http://ci.nii.ac.jp/naid/10019716933/⟩.

[10] Liam, C.: System and Method for Delivering a Human Interactive Proof to the Visually Impaired by Means of Semantic Association of Objects, USPTO Applicaton 20120232907 (2012).

[11] Michitomo, Y., Toru, N., Hajime, W., Takeshi, O. and Hiroaki, K.: Vulnerability of the Conventional Accessible CAPTCHA used by the White House and an Alternative Approach for Visually Impaired People, *Proc. 2014 IEEE International Conference on Systems, Man, and Cybernetics*, p.1729 (Paper-ID), IEEE (2014).

[12] News, B.: Blind Federation Criticises Captcha Security Test (2013), available from ⟨http://www.bbc.com/news/technology-22754006⟩.

[13] Park, G., Stuart, L.M., Taylor, U.M. and Raskin, V.: Comparing Machine and Human Abilityto Detect Phishing Emails, *Proc. 2014 IEEE International Conference on Systems, Man, and Cybernetics*, p.1956 (Paper-ID), IEEE (2014).

[14] Qvarfordt, P., Rieffel, E. and Hilbert, D.: Motion and interaction based CAPTCHA (2013). US Patent 8,601,538.

[15] Ross, S.A., Halderman, J.A. and Finkelstein, A.: Sketcha: A Captcha Based on Line Drawings of 3D Models, *Proc. 19th International Conference on World Wide Web, WWW '10*, New York, NY, USA, pp.821–830, ACM (2010).

[16] Shirali-Shahreza, S., Penn, G., Balakrishnan, R. and Ganjali, Y.: SeeSay and HearSay CAPTCHA for Mobile Interaction, *Proc. SIGCHI Conference on Human Factors in Computing Systems, CHI '13*, pp.2147–2156, ACM (2013).

[17] Shirali-Shahreza, S. and Shirali-Shahreza, M.H.: Accessibility of CAPTCHA Methods, *Proc. 4th ACM Workshop on Security and Artificial Intelligence*, pp.109–110, ACM (2011).

[18] Tam, J., Simsa, J., Hyde, S. and von Ahn, L.: Breaking Audio CAPTCHAs, *Advances in Neural Information Processing Systems 21*, pp.1625–1632, MIT Press (2008).

[19] von Ahn, L., Blum, M., Hopper, N.J. and Langford, J.: CAPTCHA: Using Hard AI Problems for Security, *Proc. EUROCRYPT*, Vol.2656, pp.294–311, Springer-Verlag (2003).

[20] Yamaguchi, M., Nakata, T., Okamoto, T. and Kikuchi, H.: An Accessible CAPTCHA System for People with Visual Disability – Generation of Human/Computer Distinguish Test with Documents on the Net, *Proc. Human-Computer Interaction International 2014*, Springer-Verlag (2014).

[21] Yamamoto, T., Tygar, J. and Nishigaki, M.: CAPTCHA Using Strangeness in Machine Translation, *2013 IEEE 27th International Conference on Advanced Information Networking and Applications*, pp.430–437 (2010).

[22] 可児潤也，鈴木徳一郎，上原章敬，山本 匠，西垣正勝：4コマ漫画CAPTCHA，情報処理学会論文誌，Vol.54, No.9, pp.2232–2243 (2013).

[23] 青空文庫：青空文庫 Aozora Bunko: 入手先 ⟨http://www.aozora.gr.jp/⟩.

[24] 山口通智：人間ロボット判別テストのバリアフリー化のためのネット上文章の採取加工技法，ヒューマンインタフェース学会論文誌，Vol.15, No.4, pp.337–352 (2013).

[25] 山口通智，岡本 健：人間ロボット判別テストのバリアフリー化のための言語的作問とその自然文生成技法，コンピュータセキュリティシンポジウム2013予稿集，pp.3D3–3, 情報処理学会 (2013).

[26] 鴨志田芳典，菊池浩明：マルコフ連鎖による合成文章の不自然さを用いたCAPTCHAの提案と安全性評価，情報処理学会論文誌，Vol.54, No.9, pp.2156–2166 (2013).

山口 通智

2003 年東北大学大学院情報科学研究科博士前期課程修了．同年（株）東芝入社．システム LSI，特にペリフェラルデバイス・システムの設計・検証業務に従事．また，2014 年筑波技術大学技術科学研究科修士課程修了後，明治大学先端数理科学研究科現象数理専攻博士後期課程に在籍．研究テーマは，情報セキュリティ技術のバアリアフリー化等．ヒューマンインタフェース学会会員．

岡本 健 （正会員）

2002 年北陸先端科学技術大学院大学博士後期課程修了．博士（情報科学）．同年東京電機大学理工学部情報科学科助手．その後，筑波大学大学院システム情報工学研究科リスク工学専攻講師を経て，2010 年筑波技術大学大学院技術科学研究科保健科学専攻准教授，現在に至る．2007 年情報処理学会論文賞．著書に『Linux ハンドブック』（オライリージャパン：共訳），『リスク工学の基礎』（コロナ社：共著）等．

菊池 浩明 （フェロー）

1988 年明治大学工学部電子通信工学科卒業．1990 年同大学院博士前期課程修了．1994 年同博士（工学）．1990 年（株）富士通研究所入社．1994 年東海大学工学部電気工学科助手．1995 年同専任講師．1999 年同助教授，2000 年同電子情報学部情報メディア学科助教授，2006 年同情報理工学部情報メディア学科教授．2008 年同情報通信学部通信ネットワーク工学科教授．1997 年カーネギーメロン大学計算機科学学部客員研究員．2013 年明治大学総合数理学部先端メディアサイエンス学科教授．WIDE プロジェクト暗号メールシステム FJPEM の開発，認証実用化実験協議会（ICAT），IPA 独創情報技術育成事業等に従事．暗号プロトコル，ネットワークセキュリティ，ファジィ論理，プライバシ保護データマイニング等に興味を持つ．1990 年日本ファジィ学会奨励賞，1993 年情報処理学会奨励賞，1996 年 SCIS 論文賞，2010 年情報処理学会 JIP Outstanding Paper Award．2013 年 IEEE AINA Best Paper Award．2014 年情報セキュリティ文化賞．電子情報通信学会，日本知能情報ファジィ学会，IEEE，ACM 各会員．情報処理学会フェロー．