

令和 6 年 6 月 19 日現在

機関番号：12103

研究種目：基盤研究(C)（一般）

研究期間：2019～2023

課題番号：19K11957

研究課題名（和文）バリアフリーを実現するCAPTCHAの開発

研究課題名（英文）Development of CAPTCHA to realize barrier free

研究代表者

岡本 健（Okamoto, Takeshi）

筑波技術大学・保健科学部・教授

研究者番号：00349797

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究では、人間の知覚に依存しないで認証を可能にするCAPTCHAを開発した。これまでは主に人間がもつ知覚能力を明示的に利用することによって、応答者が人間かコンピュータかを判別していた。一方で我々が提案するCAPTCHAでは、例えばインターネットから得られる自然文をマルコフ連鎖に基づき変形することにより問題を作成する。検証段階では応答者が「人間が本来もつ文意文脈の理解能力があるかどうか」を評価することにより認証する。これにより、利用者は視覚や聴覚といった特定知覚に依存せずに認証でき、高齢者や身体障害者でも容易にかつ安全に利用できるバリアフリー型CAPTCHAの実現を目指した。

研究成果の学術的意義や社会的意義

従来のCAPTCHAは、応答者が人間かコンピュータかを識別するために視覚や聴覚といった人間の知覚を明示的に利用してきた。これに対して、本研究では例えば文章に対して人間がもつ固有の「文意文脈を解釈する能力」を評価することにより認証する。このため加齢に伴い感覚機能が低下している高齢者や視覚や聴覚など知覚に障害がある利用者に対し、バリアフリーな認証が実現できるという点で大きな特色がある。このことは従来、人工知能や情報セキュリティの分野であったCAPTCHAが、福祉情報工学の分野にまで応用できることを示しており、こういった点で本研究の学術的意義や社会的意義は大きい。

研究成果の概要（英文）：We have developed a CAPTCHA that enables authentication without relying on human perception. Previously, the ability to differentiate between a human and a computer respondent has mainly relied on explicitly utilizing human perceptual abilities. However, our proposed CAPTCHA creates challenges by transforming natural language text obtained from the internet based on Markov chains. The verifier determines the respondent's authenticity by assessing whether they possess the inherent human ability to understand the context and meaning of the text. This approach allows for authentication without depending on specific perceptions such as vision or hearing, aiming to create a barrier-free CAPTCHA that can be easily and safely used by the elderly and people with disabilities.

研究分野：福祉情報工学、情報セキュリティ

キーワード：CAPTCHA バリアフリー型 認証技術 情報アクセシビリティ

1. 研究開始当初の背景

インターネットの普及と技術革新により、情報通信技術（ICT）が急速に発展している。これにより多くの人々が ICT を利用し、生活の質を向上させるとともに、便利さや効率の向上を実感している。一方で、高齢者や身体障害者は、感覚機能の欠損や低下などの理由により ICT が提供する多くのサービスを楽しむことができず、現代において深刻な社会問題となっている。これらの問題の一つとして、CAPTCHA（キャプチャ）と呼ばれる認証システムの不備があげられる。

現在、主流となっている CAPTCHA は、ロボットが歪んだ画像を解析しにくいという特性を利用している。しかしながら、視覚に障害を持つ人々にとってはこの方法が使えない。代替手段として、変形した音声を用いる音声型 CAPTCHA があるが、これも人間にとって認識が難しいことに加え、人間とロボットの音に関する認証能力の差が年々縮まっているため、現実的な解決策とは言えない。

このような問題を回避するために、特定の感覚に依存しない AI 問題を活用した認証方式がある。例えば、文意や文脈の解釈を問う CAPTCHA の研究では、人間とロボットの間で文意や文脈の解釈能力に差があることを利用している。例えばこれまでに、複数の文章の中から内容が関連しているものを選ばせる方法が提案されている。このような問題を作る際には、自然な文章が必要で、これは人間が作成したものでなければならない。また、安全性の重要な指標としての問題の新規性があるが、これを保証するには十分な量の文章が必要となり、インターネットで公開された文章の活用が不可欠となる。しかしながら、これらのコンテンツは公開情報であるため、攻撃者も同じ文章を検索して収集することが可能であり、識別性の要件を満たすことが難しいといった問題がある。

2. 研究の目的

本研究では、前述の問題点を解決するため、人間の知覚に依存しない CAPTCHA の実現を目指す。我々はこの方式を「バリアフリー型」の CAPTCHA と呼び、このテーマに関する研究を推し進める。従来研究では、人間とコンピュータを区別するために、主に視覚や聴覚といった特定の知覚を利用してきた。これに対して、本研究では、例えばマルコフ連鎖を用いて生成した無意味な文章に子音交替を施し、人間が文意や文脈を解釈する能力を評価する、といった方法により、特定の知覚に頼らない認証方法を提案する。

また、得られた知見や成果を醸成させることによって、より安全強度の高い CAPTCHA を構築する。本研究では主に「ワードサラダ」という機械合成文を明示的に使用し、文中のある形態素が直前の N-gram 形態素のみにより決まる連鎖型共起表現から構成する。このためワードサラダとして、「文法構造はある程度正しいが、登場する単語がある確率分布に従った」といった文を選ぶことができ、強固な安全性をもたせることができる。

3. 研究の方法

本研究では、以下のような段階を経て、安全なバリアフリー型の CAPTCHA 開発に取り組んだ。

3.1 安全性評価の理論的研究

ワードサラダ識別問題の作問アルゴリズムに関して、認証強度の観点から安全性を評価する。アルゴリズムの構成要素は、(a)公開された文章の選定、(b)類似語の使用、(c)偽の話題の挿入、の3つとなる。それぞれの要素について、情報漏えいや認証強度の脆弱性を調査する。安全性評価のために、最も厳密な指標の一つとされる、使用する文章に対する確率密度の変位差を利用する方法を採用し、サーバマシンを用いてアルゴリズムの解析と定量的な評価を行う。さらに、実用性と安全性のバランスを検証し、提案する認証システムが信頼性のある認証基盤として使用できるかどうかを評価する。

3.2 新しいアルゴリズムの提案

現在使用されている代表的なアルゴリズムを、主にマルコフ連鎖を用いたアルゴリズムに置き換えることを検討する。認識型 CAPTCHA への応用もその一環であり、暗号部分をマルコフ連鎖型アルゴリズムに置き換えて評価を行う。この変更により、実用面での通信回数、帯域幅、メモリ使用量の大幅な削減が期待できる。どの程度の削減が可能かを定量的に評価し、さらに、アルゴリズムの実用化において重要な役割を果たす電子決済サービスへの適用も検討する。これにより、従来の認証方式に対して更なる高機能化を目指す。

3.3 システムの実装

提案アルゴリズムの実装を行う。ワードサラダを生成するプログラムを作成し、前段階の結果を反映させて適切な問題が生成できるようにパラメータを調整する。このアルゴリズムは、各種計算機を使用して実装し、多角的な評価を行う。アルゴリズムの各種パラメータ選定に際しては、現在、情報漏えいによる事件が多発している現状を鑑みて、提案するアルゴリズムに高い安全性があるかどうかを再評価する。研究の進捗に応じて、専門家の協力を得て、適切な対応策を検討する。

3.4 統合化システムの構築

成果を集約し、システムを統合する。得られた成果に対して総合的な評価を行い、必要に応じてシステムの改良を進める。実用システムの実現は、実社会で求められる匿名認証の解決手段となり、その普及にも重要であると考え、重点的に取り組む。

4．研究成果

本研究で得られた主な成果を以下に示す。

4．1 機械合成文における不自然度の相対識別問題に基づく CAPTCHA

従来の音声型 CAPTCHA は、人工知能分野の文字・音声の機械認識能力の向上により、人間にとっても解答が困難となっている。特に視覚障害者にとっては、Web アクセシビリティ環境の未整備という状況も加わり、利便性が著しく低下している。本研究では、異なる不自然さを持つ機械合成文の相対比較問題を利用することで、自然文の使用を防止し、脆弱性を改善した方式を提案した。提案方式は、人間が記述した自然文と機械合成文の間にある人間が感じる違和感を定量的に評価するというアプローチにより、前述の問題点を解決している。また、実証実験を行うことで、人間による識別能力と安全性について評価した。

4．2 マルコフ連鎖を利用した新しい問題生成

これまで提案されてきた言語型 CAPTCHA の中には、無意味な単語の並びや自然文を使用している方式が主流だったが、これらは自然文の収集が難しいという点を狙った攻撃に対して脆弱だった。提案方式では、マルコフ連鎖を利用した新しい問題生成手法を採用することで、この脆弱性を克服すると同時に、公開されている文章をコーパスとして利用できる新しい機能を備えている。また、事前に十分な規模のコーパスを準備することで、問題を生成しつつ公開された文章からコーパスを収集し、マルコフ過程モデルを更新することができる。さらに、モデルを定期的に更新することで、生成される文章の多様性を保つことが可能となった。

4．3 ワードサラダを用いた言語型 CAPTCHA

従来のクイズ型 CAPTCHA に代わる新しい言語型 CAPTCHA を提案した。最初に、従来のクイズ型 CAPTCHA の安全性を分析し、その問題点を明らかにした。その結果、我々が開発した攻撃プログラムは高い確率でクイズ型 CAPTCHA を突破できることがわかった。次に、マルコフ連鎖に基づいて自然文に相当する文章を生成することで、従来の方式に存在する脆弱性を克服した。提案する方式では、ワードサラダのみを用いることで、自然文に関連する様々な問題を解決した。また、複数のワードサラダを比較して答えるという方法により、人間による正解率の低下を抑制できた。

4．4 子音交替を利用した新しい作問手法

言語型 CAPTCHA には、人間が作成した自然文を公開された文章から多量に取得し、問題を作成できるという利点がある。しかしながら、攻撃者も同じく公開された文章を検索して問題のヒントを得ることができるという欠点も内在していた。本研究では、この問題を解決するための新しい言語型 CAPTCHA を提案した。提案方法は、視覚や聴覚など特定の知覚に依存せず、方言に見られる子音交替を利用して問題文を変形させることで、人間には解読可能だが、ロボットには検索が困難になるよう工夫している。また、この手法の人間による正答率とロボットによる攻撃成功率を定量的に評価した。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 君島翔平, 岡本健
2. 発表標題 Visual SLAMを用いた視覚障害者向け自走ロボットの試作
3. 学会等名 情報処理学会 第84回全国大会
4. 発表年 2022年

1. 発表者名 芳村涼介, 外谷渉, 赤木雅弥, 岡本健, 堀江則之, 小笠原恒雄
2. 発表標題 LightGBMを用いた公開リポジトリのエクスプロイト判定手法
3. 学会等名 コンピュータセキュリティシンポジウム2023 (CSS2023)
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------