

# 文章の不自然さを利用したユーザに負担の少ない CAPTCHA方式の提案

D-19

Proposal of CAPTCHA Method  
with Less Burden on Users Using Unnaturalness of Sentences

小貫 達也 伊與田 光宏

Tatsuya ONUKI Mitsuhiro IYODA

千葉工業大学 情報工学科

Department of Computer Science, Chiba Institute of Technology

## 1. はじめに

Webサービスが普及するなかで、ボットと呼ばれる自動プログラムを使用し、不正にサービスを利用するという悪質な行為が問題となっている。このような人間とボットを識別するための方法としてCAPTCHAと呼ばれる方式が広く利用されている。しかし、主なCAPTCHA方式はすでにボットによる突破が容易であるという問題や、ノイズを大きくしすぎたり問題を難しくし過ぎた結果、人間もサービスを利用できなくなるという問題がある。

## 2. 目的

本研究では、2つの文から不自然さを感じる部分の置換を行うCAPTCHA方式を提案する。また、文章を用いたCAPTCHA方式の欠点である応答時間の長さを改善し、ユーザに使いやすいものを目指す。

## 3. 既存研究

山口ら[1]は、Webから収集した自然文をマルコフ連鎖を用いて合成文章を作成し、合成文章と自然な文章をランダムに提示し、「自然」か「不自然」かを正しく判別できるかどうかで人間とボットを判別しているCAPTCHA方式を提案した。しかし、この方式は「自然」か「不自然」かの二択なため、ランダム攻撃に対応するために複数の問題を解かなければならない。そのため文章を読む量も増え、平均応答時間は200秒近くかかるという問題がある。

## 4. 提案方式

提案方式は、文章の収集、問題の作成、CAPTCHAによる認証からなる。問題作成のための自然文はWebに公開されている文章から収集することとする。既存研究では、自然文の収集にWebの公開文章を使うと、問題として出てきた文章をWebで検索することで容易に判別ができてしまうという問題がある。しかし、提案方式では自然文をそのまま使うことはないで、Webによる検索判別することはできない。問題の作成は、データベースに格納されている自然文から2つの文章をランダムに選び、選んだ文章の一部を入れ替えることで不自然な文章をつくる。CAPTCHAによる認証では、作成した問題をユーザに提示し、不自然な箇

所を置換させる問題をいくつか出題する。これにより、回答が「自然」か「不自然」かの二択ではなくなるため、出題数を減らすことができる。正解数があらかじめ設定した閾値以上なら認証成功とする。図1に本システムの流れを示す。

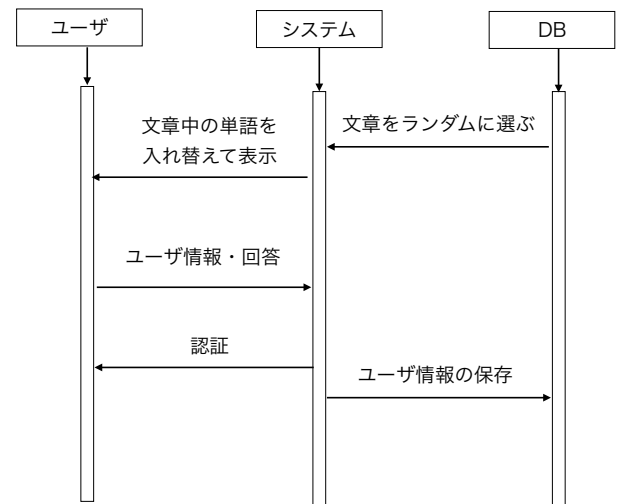


図1. CAPTCHA認証の流れ

## 5. 評価方法

提案方式を被験者に利用させ、正答率と回答時間を計測する。また、被験者にアンケートを行い、問題は難しかったかどうかを5段階で評価してもらおう。また、自由記述欄を設け、難しかったと答えた被験者には具体的にどこが難しかったかを記入してもらい、本システムの評価・考察を行う。

## 6. おわりに

本研究では、2つの文から不自然さを感じる部分の置換を行うCAPTCHA方式を提案した。これによりユーザが読む文章量が減り、ユーザにとって負担が少ない文章型のCAPTCHA方式になることを期待する。

## 参考文献

[1] 山口 通智, 岡本 健, 菊池 浩明, “機械合成文の不自然度相対識別問題に基づく CAPTCHA の提案” 情報処理学会論文誌 Vol.56 No.9 1834-1845 (Sep. 2015).