クラウド型CAPTCHAサービスにおける リスクベース認証技術の利用

荒井 毅[1] 岡部 寿男[1] 岡田 満雄[2] 渡辺 孝信[2]

概要：CAPTCHA はボットと人間を見分けるためのチューリングテストであり，CAPTCHA を用いることでボットを自動で効率的に見分けて不正アクセスを防ぐことが可能である．しかし，ボットの文字認識技術の進化により，従来型の文字型 CAPTCHA の強度ではボットによるアクセスを防ぐことができない．その一方で高度なボットを防ぐことができるような CAPTCHA では人間にとっても難しくなってしまい，利便性の低下が問題視されている．本研究では，Capy Inc. により提供されて商用サービスとして広く使われているクラウド型 CAPTCHA サービスに対してリスクベース認証技術を付加した CAPTCHA サービスの基本設計を提案する．新規のアクセスのリスクを推定し，リスクの高いアクセスに対して CAPTCHA の難度を上げるための高精度のリスク推定手法について検討した．クラウド型サービスの制約を考慮し，アクセスをユーザと紐付けられるものとユーザのわからないものに分類してそれぞれに対して異なるリスク推定手法を適用する．

1. はじめに

近年，不正アクセスによる被害の相談件数が急激に増加している．不正アクセス被害の代表的なものとしては，インターネットバンキングや仮想通貨口座からの不正送金事案が挙げられる．不正アクセスの手口としてボットを用いたパスワードの推測や総当たり法を用いた攻撃が挙げられる．従来型のユーザ ID とパスワードによる認証はボットによるアクセスによって破られやすいという弱点がある．

ボットによる自動アクセスを防ぐ手段として CAPTCHA が存在する．CAPTCHA とは，機械と人間の判別を自動で行うチューリングテストである．代表的な CAPTCHA は Google 社の運用する reCAPTCHA などが挙げられ，ボットのアクセスを防ぐ手段として広く利用されている [28]．CAPTCHA の問題生成と回答の照合は完全にカプセル化することが可能であるため，CAPTCHA はクラウド型 Web サービスとして提供することに適している [5]．その一方で，高度な光学文字認識技術を使用したボットを用いて自動的に CAPTCHA を破る手法が生み出され，認証精度の低下が問題となっている [23]．ボットの光学文字認識技術の高度化による CAPTCHA の突破の対抗手段として CAPTCHA の複雑化が進み，ユーザの利便性の低下も問題化している [7]．

CAPTCHA の複雑化に関するユーザビリティの問題に対する解決策の 1 つとして，CAPTCHA にリスクベース認証技術を CAPTCHA に応用するといった手法が考えられる．リスクベース認証では，ユーザの過去の行動や環境を分析することでアクセスのリスクレベルを判断し，リスクが高いと考えられるアクセスのみに追加の認証を課す．リスクレベルの低いアクセスに対して追加の認証を課さないことで正規ユーザの利便性を平均化するとともにリスクの高いアクセスに対する高い認証制度を実現することが可能である [12]．

本研究では，商用サービスとして現在利用されているクラウド型 CAPTCHA サービスに対してリスクベース認証技術を取り入れ，リスクの高いと判断されるアクセスに対して CAPTCHA の難度を上げるようなクラウド型 CAPTCHA システムの基本設計を提案する．

CAPTCHA の難易度をリスクベース認証のアプローチを用いて動的に変更する手法は Google 社によって開発された Invisible-reCAPTCHA によって実用化されている [19]．クラウド型の認証サービスにリスクベース認証の技術を取り入れた研究としては，Abo-alian らの研究が挙げられる [1]．Abo-alian らは，ユーザの認証時のキーストロークの癖に着目してボットによる自動入力のリスクを推定する手法を提案した．このような認証時の振る舞いの人間らしさに着目した研究は多く行われている．このような手法は

[1] 京都大学
[2] Capy 株式会社

認証精度の面で一定の成果をあげている．

本研究の主な貢献は，商用サービスとして広く利用されているクラウド型 CAPTCHA サービスに対してリスクベース認証の技術を付加し，CAPTCHA の難度を変更できるようなシステムを構築するための基本設計を示した点である．本研究では，Capy Inc. によって提供されているクラウド型 CAPTCHA サービスであるパズル CAPTCHA の過去 2 ヶ月分の認証データの 30,729,507 件を調べ，高精度のリスク推定を行う方法について検討した．提案システムにおけるリスク推定手法を用いることで，特定ユーザの普段の傾向と異なるアクセスの情報を得ることができた．また，データから得られた統計情報の分布について考察することでリスク推定のパラメータとしての有用性を示した．

以下，第 2 章では本論文の基礎となる認証技術について説明し，第 3 章では本研究に実際に用いたクラウド型 CAPTCHA について具体的な動作フローとともに説明する．第 4 章では提案するリスクベース認証技術を付加したクラウド型 CAPTCHA システムについて解説する．第 5 章においては実際にデータの分析を行い，リスク推定手法とシステムに関する考察を示す．最後に第 6 章で本論文をまとめる．

2. 準備

本節では，提案システムを解説する準備として今回用いる CAPTCHA およびリスクベース認証の技術について説明する．

2.1 CAPTCHA

CAPTCHA とは，“Completely Automated Public Turing Test To Tell Computers and Humans Apart” の略称であり，人間の手によるアクセスと機械によるアクセスを自動で判別する仕組みとして Luis von Ahn らにより 2000 年に開発された [*1]．CAPTCHA を用いて人間とボットを判別することで，人間の手を介さずにボットからのアクセスを防ぐことができる．CAPTCHA の設計概念は，ボットが解きづらく，人間に解きやすいという考えに基づいている [27]．最も一般的な画像を用いた CAPTCHA は図 1 のような画像に表示された文字列を読み取るものである．しかし，このような CAPTCHA は G.Mori らの研究 [16] によって自動的に突破が可能なことが明らかとされている．

ボットの光学文字認識（OCR）性能の進化により，ボットに破られない文字型 CAPTCHA の難易度は必然的に高くなってきている．この難易度の上昇は一部では人間にも解き難いほどになっている [3]．そのため従来の文字型 CAPTCHA 使用し続けるのではユーザビリティの低下が大きい [29]．そこで，文字型 CAPTCHA に変わる手法として，画像を用いるもの [10]，数式をもちいるもの [15]，のような人間の知識に基づく CAPTCHA の導入が提案されている．さらには，複数のノイズのような画像を重ねると文字列が浮かび上がるような手法で CAPTCHA の強度をさらに発展させるような研究も行われている [18]．

図 1: 一般的な文字列型 CAPTCHA の認証画面[*1]

2.2 リスクベース認証

本節では，リスクベース認証の仕組みについて説明する．リスクベース認証とは，ユーザが認証を行う際の環境などの情報に着目してアクセスのリスクレベルを判定し，行う認証の難しさを判断する認証方式である [2],[6],[11]．リスクレベルが低いと判断された場合，ユーザには追加の操作を要求しない．アクティブ認証と呼ばれるワンタイムパスワード認証やハードウェアトークンによる認証などに比べて認証精度は劣るが，追加の操作を要求しない点でユーザの利便性を高く保ちつつ，一定の認証制度の向上を図ることができる．

リスクベース認証で用いられる判断基準となる情報には，送信元 IP アドレスや利用 ISP 及び HTTP クライアントから User-Agent ヘッダで送信される OS やブラウザの情報，デバイスの位置情報などが主に用いられている [8],[9],[17]．IP アドレスや利用 ISP の変化の情報については，公衆無線 LAN の普及により 1 人のユーザが多数の IP アドレスと紐づけられることが増えたため，リスク推定における重要なパラメータにはならないとの研究も行われている [24]．そこで近年では，リスクベース認証の判断情報にマウス操作やキーストロークの癖のようなユーザの振る舞いに基づく情報を用いる方法も採用されている [4],[25],[26]．これらの情報がユーザの普段利用している環境の情報とどの程度異なるかを判定して数値化することで，ユーザの本人らしさを判定し，リスクレベルを判定する．リスクレベルが一定以上であると判断された場合，ユーザに追加の認証を課す．追加認証の方式はワンタイムパスワードなどの多段階認証や指紋認証のような生体情報による多要素認証などが挙げられる [21]．

リスクレベルの推定は，ユーザの過去の利用環境と今回の環境を比較して数値化することで行う．与えられたリスクレベルに対する追加認証の有無や内容については認証の重要度に応じて自由に変更することが可能である．リスクベース認証における一般的な認証フローを図 2 に示す．

[*1] The CAPTCHA Web Page: http://www.captcha.net.

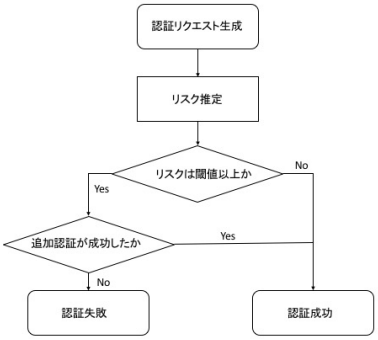

図 2: 一般的なリスクベース認証の認証フロー

3. クラウド型 CAPTCHA サービス

本節では，本研究の前提となるクラウドを利用して提供される CAPTCHA サービスに関して説明する．

3.1 クラウド型 CAPTCHA モデル

従来型の CAPTCHA では，ユーザは CAPTCHA の設置された Web サイトとの間で通信を行い回答を送信する．クラウド型 CAPTCHA では，従来型の手法とは異なり，CAPTCHA に関するシステムはクラウド上に存在し，CAPCHA の問題生成と回答の照合はクラウド上のサーバで行われる．Web サイトでは CAPTCHA の回答をクラウド上の照合サーバに送信し，照合サーバで CAPTCHA の成否が決定される．そのため，クラウド型 CAPTCHA では Web サイトとクラウド上の照合サーバの 3 者で通信を行う．モバイル端末を対象としたクラウド型の CAPTCHA サービスのスキームは A. Saxena らによって提案されている [20]．クラウド型 CAPTCHA サービスの例としては，A.Shumilov らの 3D モデルを用いた CAPTCHA などが挙げられる [22]．クラウド型 CAPTCHA サービスの具体的な通信モデルを図 3 に示す．

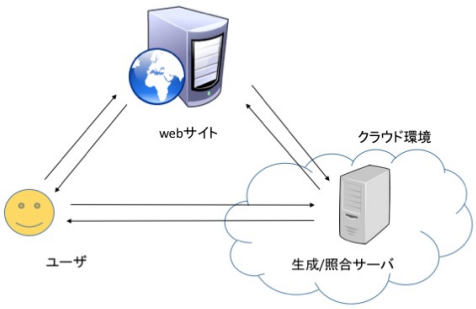

図 3: 一般的なクラウド型 CAPTCHA の通信モデル

クラウド型 CAPTCHA システムを用いる利点としては，Web サイト側でのシステム開発や運用がが不要であることによる低コスト化が挙げられる．また，クラウド型 CAPTCHA サービスの提供者は複数の Web サイトに CAPTCHA サービスを提供することで，複数の顧客から得た不審なアクセスのパターンなどの多くの CAPTCHA 情報を収集することができる．収集した情報を相関づけて分析しシステムに反映することで，個々の Web サイトがアクセスパターンを分析するよりも高い精度での CAPTCHA 情報の分析ができると考えられる．

3.2 Capy パズル CAPTCHA

パズル CAPTCHA は，Capy Inc. の提供するクラウド型 CAPTCHA サービスである[*2]．パズル CAPTCHA では一般的なクラウド型 CAPTCHA と同様に，導入する Web サイトが現在使用している認証プラットフォームにカプセル化した CAPTCHA に関するスクリプトを追加することで実装する．Web サイトの認証プラットフォームへのアクセスがあるとスクリプトにより Capy の運用する生成サーバへパズル画像の生成がリクエストされ，送信された回答は Web サイトから Capy の照合サーバへ送られて，照合サーバ上で CAPTCHA が成功しているかを判断する．Web サイト側では，CAPTCHA の成否および Web サイトの認証の成否の 2 つの要素から，ユーザの認証の成否を決定する．また、1 度 CAPTCHA に成功したユーザに関しては，次回以降 CAPTCHA を省略してログインを行うオプションを選択できる．

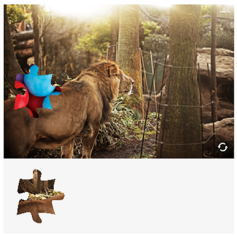

図 4: パズル CAPTCHA の認証画面
(https://www.capy.me/products/)

3.2.1 ログインの種類

パズル CAPTCHA を用いたログインの種類について説明する．パズル CAPTCHA を用いたログインの方法には，出題されたパズルに回答する puzzleCAPTCHA とパズルを省略して Cookie 値とユーザ ID の組み合わせのみでログインを行う onetimeCAPTCHA の 2 種類がある．全てのユーザは 1 度目のアクセスでは puzzleCAPTCHA を行い，puzzleCAPTCHA に成功したユーザは次回以降のログイン時に onetimeCAPTCHA を使用するかどうかの選択

[*2] https://www.capy.me/jp/products/puzzle_captcha/

を行うことができる．

3.2.1.1 puzzleCAPTCHA

パズル CAPTCHA では．ユーザは出題されたパズルに回答することで CAPTCHA の判定を行う．puzzzle 認証では，一部分が切り抜かれたパズル画像と空白部分に一致するピース画像が表示され，空白部分にピースをドラッグして重ねることで認証を行う．実行時のパズルのピースの位置が回答として送信され，照合サーバで回答の正しさを確かめることで CAPTCHA の成否が決定する．

3.2.1.2 onetimeCAPTCHA

onetimeCAPTCHA では，ユーザはユーザ ID，Cookie 値の 2 値の組み合わせを用いて CAPTCHA の判定を行う．照合サーバ内には puzzleCAPTCHA に成功したユーザ ID と Cookie 値の組み合わせが保存されている．ユーザから送信されたユーザ ID と Cookie 値の組み合わせに対して照合サーバ内でマッチするものがあれば CAPTCHA は成功となる．マッチする組み合わせがない場合 CAPTCHA は失敗となり，ユーザは再度 puzzleCAPTCHA を行うものとする．

3.2.2 通信フロー

本節では，3.2.1.1 節で挙げた 2 種類のログインの具体的な通信フローについて説明する．今後の説明の準備として，CAPTCHA を行う際にユーザが各サーバと行う通信について名前を定義する．ログインを行うユーザは，2 種類のログイン方法それぞれにおいて認証を行う Web サイトと CAPTCHA 生成/照合サーバの 2 つと通信を行っている．ユーザが Web サイトや生成サーバと行う通信を，それぞれの通信の特性により 3 つに分類し，以下の名前で定義する．

- get-js
- get-image
- verify

get-js に分類されるものは CAPTCHA に関するスクリプトおよびスクリプトの含まれるページをリクエストする通信である．get-image に分類されるものはパズル CAPTCHA で用いる画像を生成サーバにリクエストする通信である．verify に分類されるものは認証データおよび CAPTCHA の回答のポストを行う通信であるとする．

3.2.2.1 puzzleCAPTCHA の通信フロー

puzzleCAPTCHA では，ユーザによる通信は get-js，get-image，verify の順に行われる．puzzleCAPTCHA における全体の通信は図 5 で行われる．アクセスから CAPTCHA 終了までのフローは以下のようになる．

1. ユーザがパズル CAPTCHA を含む Web サイトにアクセスし，
パズル CAPTCHA 表示用の JavaScript の含まれたページをリクエストする．(get-js)
2. Web サイトがユーザにリクエストされたページを返す．
3. ユーザが生成サーバにパズル画像の表示をリクエストする．(get-image)
4. 照合サーバがユーザにリクエストされた画像を返す．
5. ユーザが Web サイトにユーザ ID，パスワード，パズルの回答をポストする．(verify)
6. Web サイトが照合サーバにパズルの回答の成否を照会する．
7. 照合サーバがパズルの回答を照合し，Web サイトに照会結果を送信する．
8. Web サイトがユーザに認証の結果に応じたページを表示する．

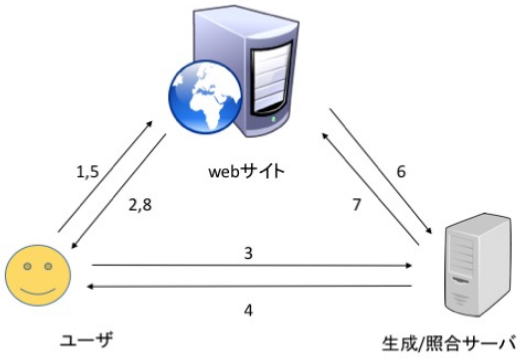

図 5: puzzleCAPTCHA の通信手順

3.2.2.2 onetimeCAPTCHA の通信フロー

onetimeCAPTCHA では，ユーザによる通信は get-js，get-js，verify の順に行われる．onetimeCAPTCHA における全体の通信は puzzleCAPTCHA と同様に図 5 のようになる．アクセスから CAPTCHA 終了までのフローについては以下のようになる．

1. ユーザがパズル CAPTCHA を含む Web サイトにアクセスし，
パズル CAPTCHA 表示用の JavaScript の含まれたページをリクエストする．(get-js)
2. Web サイトがユーザにリクエストされたページを返す．
3. ユーザが生成サーバに onetime 認証用の JavaScript をリクエストする．(get-js)
4. 照合サーバがユーザにリクエストされた JavaScript を返す．
5. ユーザが Web サイトにユーザ ID，パスワード，Cookie をポストする．(verify)
6. Web サイトが照合サーバに Cookie と ID の照合の成否を照会する．
7. 照合サーバが Cookie, ユーザ ID を照合し，Web サイトに照会結果を送信する．
8. Web サイトがユーザに認証の結果に応じたページを表示する．

3.2.3 パズル CAPTCHA の限界

本研究で利用したパズル CAPTCHA においてもボッ

トを用いて CAPTCHA を自動で破るような研究がされており，高い確率 CAPTCHA を破る成果をあげている [13],[14]．複雑なシステムの CAPTCHA を考案することはボット対策に一定の成果を示すが，ボットの進歩によって将来的に破ることが可能になると考えられるため，CAPTCHA を単体で用いたセキュリティ対策はボット対策としては不十分であると考えられる．このような問題を解決するため，CAPTCHA に対して他の認証技術のアプローチを取り入れるのが望ましいと考えられる．

4. 提案するシステム

本節では，本研究で提案するクラウド型 CAPTCHA サービスにリスクベース認証のアプローチを付加したクラウド型 CAPTCHA システムについて説明する．本研究の目標とするところは，クラウド型 CAPTCHA サービスにリスクベース認証技術を取り入れ，キャプチャの難度を高くするような仕組みを持つシステムの基本設計を示すこと，さらには高精度のリスク推定手法について検討することである．

4.1 システムの構成

提案するクラウド型 CAPTCHA システムは，既存の CAPTCHA システムの環境に対してユーザデータベースと統計分析サーバ，ユーザ分析サーバおよびリスク推定サーバを追加することで実現する．ユーザデータベースには，照合サーバから得られる認証情報をユーザごとに記録する．統計分析サーバではユーザデータベースから認証データを受け取り，IP アドレスごとの失敗率や地理情報，アクセス時間の分布について分析し，統計情報として保管する．ユーザ分析サーバにおいてはユーザデータベースから認証データを受け取り，ユーザごとのアクセス時間の傾向および利用 ISP をもとにユーザのクラスタリングを行う．リスク推定サーバにおいては，統計分析サーバ，ユーザ分析サーバからのデータをもとにアクセスのリスクを推定する．実際のシステムの構成図は図 6 のようになる．

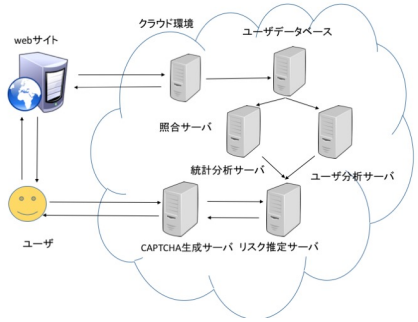

図 6: 提案システムの構成

クラウド型 CAPTCHA の問題生成にリスクベース認証のアプローチを適用した動作フローは図 7 のようになる．

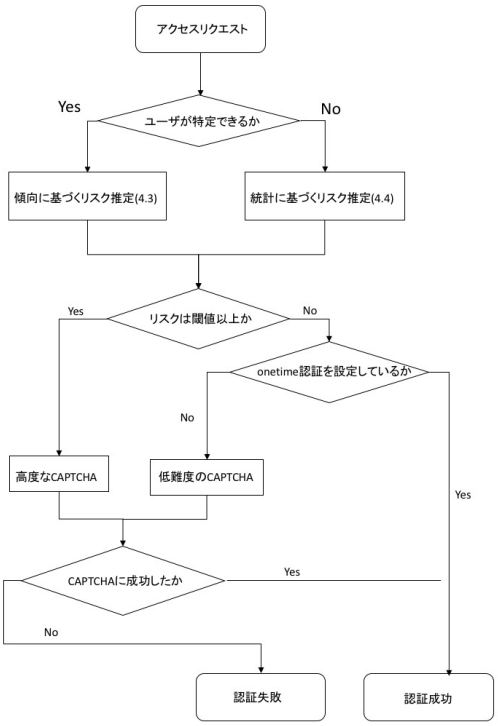

図 7: 提案システムの認証フローチャート

4.2 解析に用いる情報

CAPTCHA の難易度判断にリスクベース認証のアプローチを適用するためには，CAPTCHA の画像生成リクエスト段階でリスクレベルを推定する必要がある．ユーザから得られる情報は，get-js，get-image で得られる情報のみとなる．得られる情報はアクセス時間，アクセス元 IP アドレス，OS やブラウザの情報，Cookie 値，アクセス元 IP アドレスの利用 ISP，アクセス元の国名となる．今回はこのうちアクセス時間，アクセス元 IP アドレス，Cookie 値，利用 ISP および国名に着目してリスクレベルを推定する方法を検討する．

本システムでは，アクセスを 2 種類に分類し，それぞれ異なるリスク推定手法を用いる．アクセスがあった際に，ユーザを特定できるかを判定し，ユーザが特定できるものは 4.3 の手法を，ユーザが特定できないものは 4.4 の手法を用いてリスクの推定を行う．ユーザが特定できるかの判定には，アクセスの際の Cookie の値が照合サーバに保存されているかを用いる．

4.3 ユーザが特定できている場合の判定

本節では，ユーザが特定できるようなアクセスに対して，どのような情報を用いてリスク推定を行うことができるか検討する．

ユーザが特定できるアクセスの場合，リスクレベルの推定には従来型リスクベース認証と同様の判断基準を用いる

ことができる．推定パラメータとして，OS やブラウザの情報，アクセス時間， IP アドレスからわかるドメイン情報や利用 ISP や地理情報を利用する．準備段階で分類したユーザの傾向と比較を行い，それぞれの環境の変化についてスコア化することで本人らしさを判断し，スコアを考慮して CAPTCHA の難度を変更する．スコアが一定以下となる場合，onetime 認証を設定しているユーザに対しても CAPTCHA を課すことで認証精度を高める．

4.4 ユーザが特定できていない場合の判定

本節では，ユーザが特定できないようなアクセスに対して，どのような情報を用いてリスク推定を行うことができるか検討する．

ユーザが特定できないアクセスでは，過去の行動に基づいたリスク推定法を用いることはできない．そこで，現在のアクセス環境について，過去に蓄積された統計データと比較することでリスクの推定を行う．推定に用いるパラメータは OS やブラウザの情報，アクセス時間，IP アドレスおよび IP アドレスの逆引きで得られるドメイン名，利用 ISP および地理情報を用いる．

5. 評価と考察

本節では提案システムにおけるリスク推定手法の妥当性について，実際のパズル CAPTCHA のアクセスデータをもとに評価を行い，評価に関する考察を示す．

5.1 使用したデータセット

本節では，本研究に使用したデータセットに含まれる情報と，その詳細について述べる．本研究では，Capy Inc. から提供して頂いた，2017 年 7 月 1 日 9 時 00 分 00 秒から 2017 年 9 月 1 日 8 時 59 分 59 秒までの 2 ヶ月分のパズル CAPTCHA のアクセスログを分析用データセットとして用いる．データセットには 3.2.2 節で示した 3 種類の通信のログが蓄積されている．通信の種類ごとにログには異なる情報が記録される，通信の種類と記録される情報の対応を表 1 に示す．

表 1 に記載された 13 の情報のうち，application には，どのような認証方法を用いたかが記録される（本研究ではパズル CAPTCHA となる）．name には 3.2.2 節の 3 種類の通信のどれにあたるかが記録される．captcha_key はパズル画像 1 つ 1 つに割り振られた識別子であり，どのパズルにアクセスしているかを判別できる．challenge_key は get_image を行った際に照合サーバからユーザに割り振られる識別子であり，challenge_key を用いてユーザの 1 回分の認証データの紐付けを行う．remort_address は基本的にはエンドユーザの IP アドレスであるが，name が verify の際には remort_address は経由した Web サイトのサーバの IP アドレスとなる．よって，name の値が verify であるログには enduser_ip_address にエンドユーザの IP アドレスを別途記録している．result には通信の結果が記載される．result の詳細は付録に添付する．一般的なユーザがパズル認証を行った場合の一連のログデータの例を表 2 に示す．

表 1: 通信の種類とログの内容の対応

	get-js	get-image	verify
time	○	○	○
application	○	○	○
name	○	○	○
remort_address	○	○	○
http_user_agent	○	○	○
http_accept_language	○	○	○
captcha_key	○	○	○
challenge_key		○	○
answer			○
result	○	○	○
enduser_ip_address			○
cookie_id	○		○*1
hashed_user_id			○*1

表 2: あるユーザがパズル認証を行ったログの例 (IP アドレスは匿名化している)

(a) 時間, 通信の種類, リモートアドレス, ユーザエージェント

time	application	name	remote_address	http_user_agent
1503010659	puzzle	get_js	A.B.84.169	Mozilla/5.0 (Windows...
1503010660	puzzle	get_image	A.B.84.169	Mozilla/5.0 (Windows...
1503010706	puzzle	verify	C.D.97.31	0

(b) accept-language, パズル識別子，セッション識別子

http_accept_language	captcha_key	challenge_key
ja-JP	yLYNVsAlyIJOPq7E1X...	0
ja-JP	yLYNVsAlyIJOPq7E1X...	rGaR2s3eJhj5lvpn+...
0	yLYNVsAlyIJOPq7E1X...	rGaR2s3eJhj5lvpn+...

(c) 回答，結果，エンドユーザの IP アドレス，クッキー値，ユーザ ID

answer	result	enduser_ip_address	cookie_id	hashed_user_id
0	success	0	0	0
0	success	0	0	0
8,T,0xcgxax0x...	success	A.B.84.169	0	0

5.2 ユーザが特定できるアクセスのリスク判定法の評価

本節では，提案した手法について実際のデータセットを用いて実験を行い，得られた結果について述べる．

データセットから得られるユーザの傾向について，準備段階としてアクセス時間と IP アドレスからわかる利用 ISP の情報を用いて，ユーザを分類した．新規のアクセスについて，アクセス時間と利用 ISP が普段の傾向と異なるかを推定パラメータの 1 つとする．また，利用している OS や

*1 onetime ログインを設定しているユーザのみ記録

ブラウザの情報の変化やアクセス元の地理情報の変化も推定パラメータとして利用した．これらの推定パラメータの組み合わせから，新規のアクセスの普段の傾向との違いを判断する．

データセットの 7 月分のデータに対して分析を行い，ユーザのアクセス時間と利用 ISP の傾向を分類した．残りのデータセットにおけるユーザのアクセス時間傾向の変化，OS やブラウザの情報および利用 ISP や地理情報の変化に注目して分析を行った．本研究ではユーザを以下の 4 クラスに分類した．

(1) 利用 ISP が携帯会社のもののみであるユーザ
(2) 利用 ISP が携帯会社以外のもののみであるユーザ
(3) 9 時から 17 時は利用 ISP が携帯会社であり，他の時間は利用 ISP が携帯会社以外である割合が 80%以上のユーザ
(4) 1.2.3 のどれにも当てはまらないユーザ

分析の結果として，1,507,789 ユーザのアクセスのうち 292,017 のユーザが 7 月のアクセスデータを用いて分類した傾向とは利用時間と利用 ISP の傾向が変化していることがわかった．このようなアクセスが海外から行われたユーザは 250 人であった．具体的な過去の傾向と異なるアクセスの例を表 3 に示す．

表 3: あるユーザの普段の傾向と異なる認証ログの例

(a) ユーザ ID，時刻，利用 ISP

hashed_user_id	datetime	netname
81tfxjgAjmYixjnGC...	2017-07-31 19:47:16	KDDI
81tfxjgAjmYixjnGC...	2017-07-31 20:19:56	NTTDoCoMo
81tfxjgAjmYixjnGC...	2017-08-24 09:50:07	ZSCALERINC.ZSCAL

(b) クッキー値，結果，国コード

cookie_id	result	country
LNP1wFyLcIuPr4l...	onetime-success	JP
en9XErpj17EYnNV...	success	JP
LNP1wFyLcIu...	onetime-success	US

表 3(a) のログでは，該当ユーザは 7 月時点では携帯会社の ISP のみを利用している．しかし 3 行目のログでは携帯会社でない ISP を利用しており，さらには外国からアクセスされており，アクセス時刻も 1,2 行目と 3 行目では大きく異なっている．このようなアクセスに対しては明らかにユーザの利用環境が普段と異なるものであるため，Cookie 値が 1 度認証に使われているようなアクセスであっても onetimeCAPTCHA をせずに難度の高い CAPTCHA を課す必要がある．このようにして，提案した推定手法を用いることで，ユーザの普段の傾向と異なるアクセスを検知することが可能であると示すことができた．

5.3 ユーザが特定できないアクセスのリスク判定法の評価

ユーザが特定できないアクセスのリスク推定には，IP アドレスと OS やブラウザの情報とアクセス時間の情報を用いてリスクの推定を行う．推定の基準として，OS やブラウザの情報の変化，IP アドレスごとの失敗率，時間ごとの利用 ISP の比率とアクセス数の比率，アクセス元の国名の比率の統計情報からアクセス元の環境のリスクレベルを判定する．今回用いたログには 30,729,507 件の CAPTCHA の試行が記録されており，そのうち失敗しているものが 1,730,740 件あるため，全体の平均失敗率は約 5.6%となる．図 8 に/24 の単位で分割した IP アドレスごとの失敗率 (5%以上のもの) を示す．図 8 の失敗率は 1 の位を四捨五入したものである．図の見やすさのため，IP アドレス数の上限を 20,000 に設定しているが，10%と 20%の部分では上限以上の IP アドレスが観測されている．データに含まれていた 325,662 の/24 の IP アドレスのうち約 1.6%に相当する 5,151 件の IP アドレスで失敗率が 50%を超えている． 図 8 から，/24 単位で見た IP アドレスについて成功率下位 1%に入るようなものは他のものと比べて十分に失敗率が高いことがわかる．このことから本データセットにおいては IP アドレスは統計的リスク推定において十分に利用可能なパラメータであると考えられる．

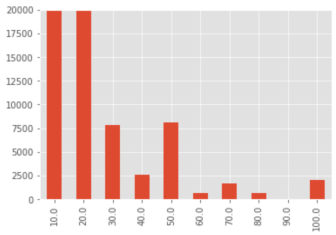

図 8: IP アドレスごとの失敗率 (横軸:失敗率 [%], 縦軸:該当 IP アドレス数)

次に，アクセス元の国名に注目する． 国別のアクセス数上位 30 国の国コード (ISO3166-1 alpha-2) を表 4 に示す．国別のアクセス数は 98.5%が日本からのアクセスである．次点がアメリカからのアクセスとなり，全体の 0.8%となる．このことから本データセットにおいてはアクセス元の国が日本であるかどうかが統計的リスクの推定に利用可能なパラメータであることがわかる．

次にユーザのアクセス時間の傾向に注目する．時間ごとのアクセス数と携帯会社の回線を利用したアクセスの比率は図 9，図 10 のようになる．どちらのデータにも，時間ごとに一定の変化が見られる．携帯会社の回線の利用率について，12 時台の利用が他の時刻と比べて高くなっている．このような傾向は昼休みの時間にスマートフォンを利用したアクセスが多いためと考えられる．このような利用回線の比率やアクセス時間についても推定のパラメータと

表 4: 国別のアクセス数上位 30 件 (ISO3166-1 alpha-2)

国コード	件数	国コード	件数
JP	30,278,096	PH	2,553
US	255,821	MY	2,500
TH	9,420	IN	2,401
CN	7,973	MX	1,430
SG	6,901	IT	1,334
GB	5,990	private	1,196
HK	5,289	NZ	1,067
AU	4,164	CH	1,049
ID	4,099	NL	1,041
DE	4,077	BE	884
VN	3,975	MM	835
CA	3,963	BR	819
KR	3,694	SE	805
TW	3,448	ES	802
FR	2,994	KH	528

して利用できると考えられる．

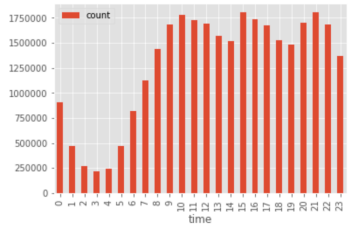

図 9: 時間帯ごとのアクセスの数 (横軸：時刻，縦軸：アクセス回数)

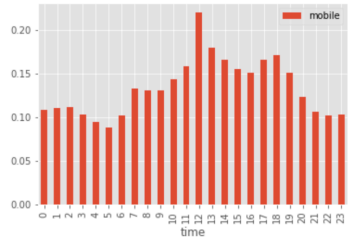

図 10: 時間帯ごとのモバイル回線を利用したアクセスの比率 (横軸：時刻，縦軸：比率 [%])

5.4 考察

本節では，提案手法と実際のデータから得られた実験結果による考察を示す．提案システムで用いた 2 種類のリスク推定手法について，それぞれの手法に基づく推定で一定の結果を得ることができた．ユーザの特定できるアクセスについてはユーザの傾向を事前に分類しておくことで，普段の傾向と異なるアクセスログを発見することができた．ユーザの特定できないアクセスについて，IP アドレスごとの CAPTCHA 失敗率の高さを中心にリスクの高い環境についての統計的データを得ることができた．

本研究では，手動でデータの分析を行うことでリスク推定手法の妥当性を示した．本手法で示した行動分析およびリスク推定手法を自動化し，一定のリスク分析の成果が得られることを示すことは本研究の大きな 1 つの課題である．また，クラウド型サービスの利点である複数の Web サイトの統計データを相関づけて分析できる点に対して，実際に各 Web サイトの統計データからリスク推定の助けとなる相関が得られるかを検証することも課題の 1 つである．本研究の結果を有効に活用するために，CAPTCHA の難易度をどのように判定して分類するのかの手法についても考慮しなければならない．本研究の発展として，GeoIP[*3]の有料サービスを利用して国だけでなく都市単位で IP アドレスの分類を行うことで，ユーザの行動傾向をさらに正確に読み取ることができると考えられる．

6. まとめ

本研究では，クラウド型 CAPTCHA サービスのユーザの利便性の向上および認証精度の向上を目的として，リスクベース認証で用いられるリスク推定技術の適用方法を提案した．Cookie 値を用いてアクセスを 2 種類に分類し，それぞれ異なる手法を用いてリスクを推定することで，より正確性の高いリスク推定を行うことを検討した．

クラウド型 CAPTCHA サービスの CAPTCHA 生成にリスクベース認証技術を適用するためには，全てのアクセスに対してユーザを特定することができないという点で制約が存在する．この制約に対処するため，ユーザの特定できるものとそうでないものそれぞれについてリスク推定手法を検討した．ユーザの特定できるアクセスについては従来型のリスクベース認証と同様に過去の傾向に基づいたリスク評価を行う手法について検討した．ユーザの特定できないアクセスについては，IP アドレスや地理情報の傾向から統計データを作成し，アクセスの環境と比較することでリスク推定を行う手法について検討した．このようなリスク推定方法を用いることで，ユーザが特定できるアクセスに対して，過去の環境との比較から傾向の異なるアクセスを見つけることができた．また，ユーザが特定できないアクセスに対しても統計データからのリスク推定を行うことは十分可能であると示すことができた．

本研究の今後の課題として，今回手動でユーザの行動分析を行ったものについて，より正確にユーザの行動パターンを定義して自動化することで分析の質を高める必要がある．また，複数 Web サイトから得られる統計データから相関を得て，統計的リスク推定の質を高めることができるかについても調査の必要がある．また，IP アドレスから都市の情

[*3] https://www.maxmind.com/ja/home

報を得ることでさらに行動分析を発展させることができると考えられる．また，本研究の CAPTCHA は JavaScript で記述されているため，リスク推定を CAPTCHA 完了時に設定することでユーザの回答の際の挙動を用いたリスクベース認証技術の適用を行うことも可能である．

参考文献

[1] Abo-alian, A., Badr, N. L. and Tolba, M. F.: Authentication As a Service for Cloud Computing, *Proceedings of the International Conference on Internet of Things and Cloud Computing*, ICC '16, New York, NY, USA, ACM, pp. 10:1–10:7 (2016).

[2] Bakar, K. A. A. and Haron, G. R.: Adaptive authentication: Issues and challenges, *2013 World Congress on Computer and Information Technology (WCCIT)*, pp. 1–6 (2013).

[3] Beheshti, S. M. R. S. and Liatsis, P.: CAPTCHA Usability and Performance, How to Measure the Usability Level of Human Interactive Applications Quantitatively and Qualitatively?, *2015 International Conference on Developments of E-Systems Engineering (DeSE)*, pp. 131–136 (2015).

[4] Chakraborty, A., Munshi, S. and Kundu, A.: An Adaptive Server Side Software Authentication Framework Based on User's Activity Pattern, *2011 Second International Conference on Emerging Applications of Information Technology*, pp. 153–156 (2011).

[5] Converse, T.: CAPTCHA Generation as a Web Service, *Human Interactive Proofs* (Baird, H. S. and Lopresti, D. P.(eds.)), Berlin, Heidelberg, Springer Berlin Heidelberg, pp. 82–96 (2005).

[6] Diep, N., Lee, S., Lee, Y. and Lee, H.: *Contextual risk-based access control*, pp. 406–412 (2007).

[7] Fidas, C. A., Voyiatzis, A. G. and Avouris, N. M.: On the Necessity of User-friendly CAPTCHA, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, New York, NY, USA, ACM, pp. 2623–2626 (2011).

[8] Fridman, L., Weber, S., Greenstadt, R. and Kam, M.: Active Authentication on Mobile Devices via Stylometry, Application Usage, Web Browsing, and GPS Location, *IEEE Systems Journal*, Vol. 11, No. 2, pp. 513–521 (2017).

[9] G., T.: Emergence of Risk-Based Authentication in Online Financial Services: You Can't Hide Your Lyin' IPs, Whitepaper v43:15n, TowerGroup (2005).

[10] Gao, H., Yao, D., Liu, H., Liu, X. and Wang, L.: A Novel Image Based CAPTCHA Using Jigsaw Puzzle, *2010 13th IEEE International Conference on Computational Science and Engineering*, pp. 351–356 (2010).

[11] Golan, L., Orad, A. and Bennett, N.: System and method for risk based authentication (2005).

[12] Han, W., Sun, C., Shen, C., Lei, C. and Shen, S.: Dynamic combination of authentication factors based on quantified risk and benefit, *Security and Communication Networks*, Vol. 7, No. 2, pp. 385–396 (2014).

[13] Hernåndez-Castro, C. J., R-Moreno, M. D. and Barrero, D. F.: Side-Channel Attack against the Capy HIP, *2014 Fifth International Conference on Emerging Security Technologies*, pp. 99–104 (2014).

[14] Hernåndez-Castro, C. J., R-Moreno, M. D. and Barrero, D. F.: Using JPEG to Measure Image Continuity and Break Capy and Other Puzzle CAPTCHAs, *IEEE Internet Computing*, Vol. 19, No. 6, pp. 46–53 (2015).

[15] Hernandez-Castro, C. J. and Ribagorda, A.: Pitfalls in CAPTCHA design and implementation: The Math CAPTCHA, a case study, *Computers & Security*, Vol. 29, No. 1, pp. 141 – 157 (2010).

[16] Mori, G. and Malik, J.: Recognizing objects in adversarial clutter: breaking a visual CAPTCHA, *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, Vol. 1, pp. I–134–I–141 vol.1 (2003).

[17] Obaidat, M. S. and Macchiarolo, D. T.: An online neural network system for computer access security, *IEEE Transactions on Industrial Electronics*, Vol. 40, No. 2, pp. 235–242 (1993).

[18] Okada, M. and Matsuyama, S.: New CAPTCHA for smartphones and tablet PC, *2012 IEEE Consumer Communications and Networking Conference (CCNC)*, pp. 34–35 (2012).

[19] Powell, B. M., Singh, R., Vatsa, M. and Noore, A.: Poster: Adaptcha: An Adaptive CAPTCHA for Improved User Experience, *system*, Vol. 4, p. 6.

[20] Saxena, A., Chauhan, N. S., Reddy, S. K., Vangal, A. S. and Rodrguez, D. P.: A New Scheme for Mobile Based CAPTCHA Service on Cloud, *2012 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM)*, pp. 1–6 (2012).

[21] Shah, Y., Choyi, V. and Subramanian, L.: Multi-factor Authentication as a Service, *2015 3rd IEEE International Conference on Mobile Cloud Computing, Services, and Engineering*, pp. 144–150 (2015).

[22] Shumilov, A. and Philippovich, A.: Cloud-based CAPTCHA service, *2016 6th International Conference - Cloud System and Big Data Engineering (Confluence)*, pp. 115–118 (2016).

[23] Sivakorn, S., Polakis, J. and Keromytis, A. D.: I' m not a human: Breaking the Google reCAPTCHA, *Black Hat* (2016).

[24] Spooren, J., Preuveneers, D. and Joosen, W.: Mobile Device Fingerprinting Considered Harmful for Risk-based Authentication, *Proceedings of the Eighth European Workshop on System Security*, EuroSec '15, New York, NY, USA, ACM, pp. 6:1–6:6 (2015).

[25] Traore, I., Woungang, I., Obaidat, M. S., Nakkabi, Y. and Lai, I.: Combining Mouse and Keystroke Dynamics Biometrics for Risk-Based Authentication in Web Environments, *2012 Fourth International Conference on Digital Home*, pp. 138–145 (2012).

[26] Traore, I., Woungang, I., Obaidat, M. S., Nakkabi, Y. and Lai, I.: Online risk-based authentication using behavioral biometrics, *Multimedia Tools and Applications*, Vol. 71, No. 2, pp. 575–605 (2014).

[27] von Ahn, L., Blum, M. and Langford, J.: Telling Humans and Computers Apart Automatically, *Commun. ACM*, Vol. 47, No. 2, pp. 56–60 (2004).

[28] von Ahn, L., Maurer, B., McMillen, C., Abraham, D. and Blum, M.: reCAPTCHA: Human-Based Character Recognition via Web Security Measures, *Science*, Vol. 321, No. 5895, pp. 1465–1468 (2008).

[29] Yan, J. and El Ahmad, A. S.: Usability of CAPTCHAs or Usability Issues in CAPTCHA Design, *Proceedings of the 4th Symposium on Usable Privacy and Security*, SOUPS '08, New York, NY, USA, ACM, pp. 44–52 (2008).