

(19)日本国特許庁(JP)

(12)特許公報(B1)

(11)特許番号
特許第7176158号
(P7176158)

(45)発行日 令和4年11月21日(2022.11.21)

(24)登録日 令和4年11月11日(2022.11.11)

(51)国際特許分類	F I
G 0 6 F 21/55 (2013.01)	G 0 6 F 21/55
G 0 6 N 20/00 (2019.01)	G 0 6 N 20/00

請求項の数 15 (全43頁)

(21)出願番号	特願2022-529394(P2022-529394)	(73)特許権者	399037405 楽天グループ株式会社 東京都世田谷区玉川一丁目14番1号
(86)(22)出願日	令和3年6月30日(2021.6.30)	(74)代理人	110000154弁理士法人はるか国際特許事務所
(86)国際出願番号	PCT/JP2021/024841	(72)発明者	友田 恭輔 東京都世田谷区玉川一丁目14番1号 楽天グループ株式会社内
審査請求日	令和4年5月19日(2022.5.19)	(72)発明者	伊藤 周平 東京都港区港南二丁目16番5号 NB F品川タワー 楽天ペイメント株式会社内
早期審査対象出願		審査官	岸野 徹
最終頁に続く			

(54)【発明の名称】 学習モデル評価システム、学習モデル評価方法、及びプログラム

(57)【特許請求の範囲】

【請求項1】

所定のサービスを利用可能なユーザ端末から前記サービスにログインした状態で所定の認証を実行した認証済みユーザの行動に関する認証済み情報を取得する認証済み情報取得手段と、

前記認証済み情報に基づいて、前記サービスにおける不正を検知するための学習モデルからの出力を取得する出力取得手段と、

前記認証済み情報に対応する前記出力に基づいて、前記学習モデルの精度を評価する評価手段と、

を含む学習モデル評価システム。

【請求項2】

所定のサービスを利用可能なユーザ端末を利用して、前記サービスで利用可能な第1カードを所持しているか否かを確認するための所持認証を実行した認証済みユーザの前記第1カードに対応する行動に関する認証済み情報を取得する認証済み情報取得手段と、

前記第1カードに対応する前記認証済み情報に基づいて、前記サービスにおける不正を検知するための学習モデルからの前記第1カードに対応する出力を取得する出力取得手段と、

前記第1カードに対応する前記出力に基づいて、前記学習モデルの精度を評価する評価手段と、

前記第1カードの名義に関する第1名義情報と、前記認証済みユーザが前記サービスで

利用可能な第2カードの名義に関する第2名義情報と、を比較する比較手段と、
を含み、

前記認証済み情報取得手段は、前記比較手段の比較結果が所定の結果である場合に、前記第2カードに対応する前記認証済み情報を取得し、

前記出力取得手段は、前記第2カードに対応する前記認証済み情報に基づいて、前記第2カードに対応する前記出力を取得し、

前記評価手段は、前記第1カードに対応する前記出力と、前記第2カードに対応する前記出力と、に基づいて、前記学習モデルの精度を評価する、

学習モデル評価システム。

【請求項3】

前記認証済み情報取得手段は、複数の前記認証済み情報を取得し、

前記出力取得手段は、前記複数の認証済み情報の各々に対応する前記出力を取得し、

前記評価手段は、前記複数の認証済み情報の各々に対応する前記出力に基づいて、前記学習モデルの精度を評価する、

請求項1又は2に記載の学習モデル評価システム。

【請求項4】

前記学習モデル評価システムは、前記学習モデルの精度が所定の精度未満になった場合に、前記サービスにおける最近の行動を利用して前記学習モデルを作成するための処理を実行する処理実行手段、

を更に含む請求項1～3の何れかに記載の学習モデル評価システム。

【請求項5】

前記学習モデル評価システムは、不正であるか否かが確定した確定ユーザの行動に関する確定情報を取得する確定情報取得手段を更に含み、

前記評価手段は、前記認証済み情報及び前記確定情報に基づいて、前記学習モデルの精度を評価する、

請求項1～4の何れかに記載の学習モデル評価システム。

【請求項6】

前記第2カードは、前記所持認証に対応していないカードであり、

前記第2カードに対応する前記認証済み情報は、前記所持認証が実行されていない前記第2カードを利用した前記認証済みユーザの行動に関する情報である、

請求項2に記載の学習モデル評価システム。

【請求項7】

前記学習モデル評価システムは、前記認証済み情報に基づいて、前記認証済みユーザの行動が正当と推定されるように、前記サービスにおける不正を検知するための前記学習モデルを作成する作成手段を更に含む、

請求項1～6の何れかに記載の学習モデル評価システム。

【請求項8】

前記学習モデルは、教師有り学習のモデルであり、

前記作成手段は、前記認証済み情報に基づいて、前記認証済みユーザの行動が正当であることを示す第1訓練データを作成し、前記第1訓練データに基づいて、前記学習モデルを学習させることによって、前記学習モデルを作成する、

請求項7に記載の学習モデル評価システム。

【請求項9】

前記学習モデル評価システムは、未認証ユーザの行動に関する未認証情報を取得する未認証情報取得手段を更に含み、

前記作成手段は、前記未認証情報に基づいて、前記未認証ユーザの行動が正当又は不正であることを示す第2訓練データを作成し、前記第2訓練データに基づいて、前記学習モデルを学習させる、

請求項8に記載の学習モデル評価システム。

【請求項10】

10

20

30

40

50

前記作成手段は、前記未認証情報に基づいて、学習済みの前記学習モデルからの出力を取得し、当該出力に基づいて、前記第2訓練データを作成する、

請求項9に記載の学習モデル評価システム。

【請求項11】

前記サービスは、前記ユーザ端末から利用可能な電子決済サービスであり、

前記認証済み情報は、前記電子決済サービスにおける前記認証済みユーザの行動に関する情報であり、

前記学習モデルは、前記電子決済サービスにおける不正を検知するためのモデルである、

請求項1～10の何れかに記載の学習モデル評価システム。

【請求項12】

コンピュータが、

所定のサービスを利用可能なユーザ端末から前記サービスにログインした状態で所定の認証を実行した認証済みユーザの行動に関する認証済み情報を取得する認証済み情報取得ステップと、

前記認証済み情報に基づいて、前記サービスにおける不正を検知するための学習モデルからの出力を取得する出力取得ステップと、

前記認証済み情報に対応する前記出力に基づいて、前記学習モデルの精度を評価する評価ステップと、

を実行する学習モデル評価方法。

【請求項13】

コンピュータが、

所定のサービスを利用可能なユーザ端末を利用して、前記サービスで利用可能な第1カードを所持しているか否かを確認するための所持認証を実行した認証済みユーザの前記第1カードに対応する行動に関する認証済み情報を取得する認証済み情報取得ステップと、

前記第1カードに対応する前記認証済み情報に基づいて、前記サービスにおける不正を検知するための学習モデルからの前記第1カードに対応する出力を取得する出力取得ステップと、

前記第1カードに対応する前記出力に基づいて、前記学習モデルの精度を評価する評価ステップと、

前記第1カードの名義に関する第1名義情報と、前記認証済みユーザが前記サービスで利用可能な第2カードの名義に関する第2名義情報と、を比較する比較ステップと、

を実行し、

前記認証済み情報取得ステップは、前記比較ステップの比較結果が所定の結果である場合に、前記第2カードに対応する前記認証済み情報を取得し、

前記出力取得ステップは、前記第2カードに対応する前記認証済み情報に基づいて、前記第2カードに対応する前記出力を取得し、

前記評価ステップは、前記第1カードに対応する前記出力と、前記第2カードに対応する前記出力と、に基づいて、前記学習モデルの精度を評価する、

学習モデル評価方法。

【請求項14】

所定のサービスを利用可能なユーザ端末から前記サービスにログインした状態で所定の認証を実行した認証済みユーザの行動に関する認証済み情報を取得する認証済み情報取得手段、

前記認証済み情報に基づいて、前記サービスにおける不正を検知するための学習モデルからの出力を取得する出力取得手段、

前記認証済み情報に対応する前記出力に基づいて、前記学習モデルの精度を評価する評価手段、

としてコンピュータを機能させるためのプログラム。

【請求項15】

所定のサービスを利用可能なユーザ端末を利用して、前記サービスで利用可能な第1カ

10

20

30

40

50

ードを所持しているか否かを確認するための所持認証を実行した認証済みユーザの前記第1カードに対応する行動に関する認証済み情報を取得する認証済み情報取得手段、

前記第1カードに対応する前記認証済み情報に基づいて、前記サービスにおける不正を検知するための学習モデルからの前記第1カードに対応する出力を取得する出力取得手段、

前記第1カードに対応する前記出力に基づいて、前記学習モデルの精度を評価する評価手段、

前記第1カードの名義に関する第1名義情報と、前記認証済みユーザが前記サービスで利用可能な第2カードの名義に関する第2名義情報と、を比較する比較手段、

としてコンピュータを機能させ、

前記認証済み情報取得手段は、前記比較手段の比較結果が所定の結果である場合に、前記第2カードに対応する前記認証済み情報を取得し、

前記出力取得手段は、前記第2カードに対応する前記認証済み情報に基づいて、前記第2カードに対応する前記出力を取得し、

前記評価手段は、前記第1カードに対応する前記出力と、前記第2カードに対応する前記出力と、に基づいて、前記学習モデルの精度を評価する、

プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、学習モデル評価システム、学習モデル評価方法、及びプログラムに関する。

【背景技術】

【0002】

従来、所定のサービスを利用するユーザの不正を検知する技術が知られている。例えば、特許文献1には、ユーザの行動に関する特徴量を入力とし、この行動が不正であるか否かを出力とする訓練データを、教師有り学習を利用した学習モデルに学習させることによって、サービスにおける不正を検知するための学習モデルを作成するシステムが記載されている。

【先行技術文献】

【特許文献】

【0003】

【文献】国際公開第2019/049210号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

しかしながら、サービスにおけるユーザの行動は日々変化するので、最近の傾向が学習モデルに学習されていなければ、特許文献1のような学習モデルの精度が徐々に落ちることがある。例えば、学習モデルの不正検知の精度が落ちると、本当は不正な行動だったとしても、正当と推定される可能性がある。逆に、本当は正当な行動だったとしても、不正と推定される可能性がある。このため、学習モデルの不正検知の精度を正確に評価することは重要である。

【0005】

本開示の目的の1つは、サービスにおける不正を検知するための学習モデルの精度を正確に評価することである。

【課題を解決するための手段】

【0006】

本開示の一態様に係る学習モデル評価システムは、所定のサービスを利用可能なユーザ端末から所定の認証を実行した認証済みユーザの行動に関する認証済み情報を取得する認証済み情報取得手段と、前記認証済み情報に基づいて、前記サービスにおける不正を検知するための学習モデルからの出力を取得する出力取得手段と、前記認証済み情報に対応する前記出力に基づいて、前記学習モデルの精度を評価する評価手段と、を含む。

【発明の効果】

【0007】

本開示によれば、サービスにおける不正を検知するための学習モデルの精度を正確に評価できる。

【図面の簡単な説明】

【0008】

【図1】不正検知システムの全体構成の一例を示す図である。

【図2】利用登録の流れの一例を示す図である。

【図3】所持認証の流れの一例を示す図である。

【図4】カードのICチップをNFC部で読み取る様子の一例を示す図である。

10

【図5】学習モデルの一例を示す図である。

【図6】第1実施形態の不正検知システムで実現される機能の一例を示す機能ブロック図である。

【図7】ユーザデータベースのデータ格納例を示す図である。

【図8】訓練データベースのデータ格納例を示す図である。

【図9】第1実施形態において実行される処理の一例を示すフロー図である。

【図10】第2実施形態の概要を示す図である。

【図11】第2実施形態の不正検知システムで実現される機能の一例を示す機能ブロック図である。

【図12】第2実施形態において実行される処理の一例を示すフロー図である。

20

【図13】変形例1-1の不正検知システムの全体構成の一例を示す図である。

【図14】変形例1-1のユーザ端末に表示される画面の一例を示す図である。

【図15】カードの登録後に上限額を増やす流れの一例を示す図である。

【図16】カードのICチップをNFC部で読み取る様子の一例を示す図である。

【図17】第1実施形態に係る変形例における機能ブロック図である。

【図18】ユーザデータベースのデータ格納例を示す図である。

【図19】第2実施形態に係る変形例における機能ブロック図

【発明を実施するための形態】

【0009】

[1. 第1実施形態]

30

以降、本開示に係る学習モデル作成システムの実施形態の一例である第1実施形態を説明する。第1実施形態では、学習モデル作成システムを不正検知システムに適用した場合を例に挙げる。このため、第1実施形態で不正検知システムと記載した箇所は、学習モデル作成システムと読み替えることができる。学習モデル作成システムは、学習モデルの作成までを行い、不正検知自体は、他のシステムで実行されてもよい。即ち、学習モデル作成システムは、不正検知システムのうちの不正検知の機能を含まなくてもよい。

【0010】

[1-1. 不正検知システムの全体構成]

図1は、不正検知システムの全体構成の一例を示す図である。図1に示すように、不正検知システムSは、サーバ10及びユーザ端末20を含む。サーバ10及びユーザ端末20の各々は、インターネット等のネットワークNに接続可能である。不正検知システムSは、少なくとも1つのコンピュータを含めばよく、図1の例に限られない。例えば、サーバ10は、複数台存在してもよい。ユーザ端末20は、1台だけであってもよいし、3台以上存在してもよい。

40

【0011】

サーバ10は、サーバコンピュータである。サーバ10は、制御部11、記憶部12、及び通信部13を含む。制御部11は、少なくとも1つのプロセッサを含む。記憶部12は、RAM等の揮発性メモリと、ハードディスク等の不揮発性メモリと、を含む。通信部13は、有線通信用の通信インタフェースと、無線通信用の通信インタフェースと、の少なくとも一方を含む。

50

【0012】

ユーザ端末20は、ユーザのコンピュータである。例えば、ユーザ端末20は、スマートフォン、タブレット端末、ウェアラブル端末、又はパーソナルコンピュータである。ユーザ端末20は、制御部21、記憶部22、通信部23、操作部24、表示部25、撮影部26、及びICチップ27を含む。制御部21及び記憶部22の物理的構成は、それぞれ制御部11及び記憶部12と同様である。

【0013】

通信部23の物理的構成は、通信部13と同様であってもよいが、第1実施形態の通信部23は、更にNFC（Near field communication）部23Aを含む。NFC部23Aは、NFC用の通信インタフェースを含む。NFC自体は、種々の規格を利用可能であり、例えば、ISO/IEC18092又はISO/IEC21481といった国際標準規格を利用可能である。NFC部23Aは、規格に準じたアンテナ等のハードウェアを含み、例えば、リーダ／ライタ機能、ピアツーピア機能、カードエミュレーション機能、ワイヤレス充電機能、又はこれらの組み合わせを実現する。

10

【0014】

操作部24は、タッチパネル等の入力デバイスである。表示部25は、液晶ディスプレイ又は有機ELディスプレイである。撮影部26は、少なくとも1台のカメラを含む。ICチップ27は、NFCに対応したチップである。ICチップ27は、任意の規格のチップであってよく、例えば、Felica（登録商標）のチップ、又は、非接触型規格におけるいわゆるType A若しくはType Bのチップである。ICチップ27は、規格に応じたアンテナ等のハードウェアを含み、例えば、ユーザが利用するサービスに必要な情報を記憶する。

20

【0015】

なお、記憶部12、22に記憶されるプログラム及びデータの少なくとも一方は、ネットワークNを介して供給されてもよい。また、サーバ10及びユーザ端末20の少なくとも一方に、コンピュータ読み取り可能な情報記憶媒体を読み取る読取部（例えば、光ディスクドライブやメモリカードスロット）と、外部機器とデータの入出力をするための入出力部（例えば、USBポート）と、の少なくとも一方が含まれてもよい。例えば、情報記憶媒体に記憶されたプログラム及びデータの少なくとも一方が、読取部及び入出力部の少なくとも一方を介して供給されてもよい。

30

【0016】

[1-2. 第1実施形態の概要]

不正検知システムSは、ユーザに提供されるサービスにおける不正を検知する。不正とは、違法行為、サービスの利用規約に違反する行為、又はその他の迷惑行為である。本実施形態では、他人のユーザID及びパスワードでログインし、他人になりすましてサービスを利用する行為が不正に相当する場合を例に挙げる。このため、この行為について説明している箇所は、不正と読み替えることができる。不正検知システムSは、種々の不正を検知可能である。他の不正の例は、後述の変形例で説明する。

【0017】

不正を検知するとは、不正の有無を推定又は判定することである。例えば、不正であるか否かを示す情報を出力すること、又は、不正の疑いの高さを示すスコアを出力することは、不正を検知することに相当する。例えば、スコアが数値で表現される場合、スコアが高いほど不正の疑いが高い。スコアは、数値以外にもSランク、Aランク、Bランクといったような文字等で表現されてもよい。スコアは、不正の確率又は蓋然性ということもできる。

40

【0018】

第1実施形態では、サービスの一例として、官公庁等の公的機関が提供する行政サービスを挙げる。他のサービスの例は、変形例で説明する。第1実施形態では、行政サービスを、単にサービスと記載する。第1実施形態では、サーバ10が、サービスの提供と、不正の検知と、の両方を行う場合を説明するが、サーバ10以外のコンピュータがサービス

50

を提供してもよい。ユーザ端末 20 には、公的機関のアプリケーション（以降、単にアプリ）がインストールされている。ユーザは、初めてサービスを利用する場合、サービスへのログインに必要なユーザ ID を発行するために、サービスの利用登録を行う。

【0019】

図 2 は、利用登録の流れの一例を示す図である。図 2 に示すように、ユーザがユーザ端末 20 のアプリを起動すると、利用登録に必要な情報を入力するための登録画面 G 1 が表示部 25 に表示される。例えば、ユーザは、入力フォーム F 10 に、希望するユーザ ID、パスワード、氏名、住所、電話番号、及びユーザの個人番号といった情報を入力する。ユーザ ID は、サービスにおいてユーザを一意的に識別可能な情報である。個人番号は、公的機関が発行した個人番号カードに記載された個人を識別可能な情報である。第 1 実施形態では、個人番号カードを、単にカードと記載する。

10

【0020】

ユーザがボタン B 11 を選択すると、入力フォーム F 10 に入力された情報がサーバ 10 に送信され、利用登録が完了したことを示す完了画面 G 2 が表示部 25 に表示される。利用登録が完了すると、ユーザは、アプリからサービスを利用できるようになる。例えば、ユーザがボタン B 20 を選択すると、アプリのトップ画面 G 3 が表示部 25 に表示される。例えば、トップ画面 G 3 には、アプリから利用可能なサービスの一覧が表示される。例えば、ユーザがボタン B 30 を選択すると、証明書の請求や窓口の予約といったサービスを利用するための利用画面 G 4 が表示部 25 に表示される。

【0021】

20

第三者は、フィッシング等により、ユーザ ID 及びパスワードを不正に入手することがある。この場合、第三者は、他人になりすましてサービスにログインし、サービスを不正に利用する可能性がある。そこで、第 1 実施形態では、第三者による不正利用を抑制するために、カードを利用した所持認証が実行されるようになっている。所持認証は、正当な者だけが所持する所持物を利用した認証である。所持物は、カードに限られず、任意の物であってよい。例えば、所持物は、情報記憶媒体又は用紙であってもよい。所持物は、有体物に限られず、電子的なデータのような無体物であってもよい。

【0022】

所持認証を実行するか否かは、ユーザの任意である。ユーザは、所持認証を実行せずにサービスを利用することもできる。ただし、所持認証を実行していない状態では、ユーザが利用可能なサービスが制限されている。ユーザが、自身のユーザ端末 20 から所持認証を実行すると、このユーザ端末 20 から利用可能なサービスの種類が増える。ただし、所持認証を実行したユーザのユーザ ID で他のユーザ端末 20 からログインしても、当該他のユーザ端末 20 で所持認証が実行されていなければ、当該他のユーザ端末 20 から利用可能なサービスは制限される。

30

【0023】

図 3 は、所持認証の流れの一例を示す図である。図 2 のトップ画面 G 3 のボタン B 31 が選択されると、図 3 に示すように、所持認証を開始するための開始画面 G 5 が表示部 25 に表示される。第 1 実施形態では、所持認証として、NFC を利用した NFC 認証と、画像を利用した画像認証と、の 2 種類が用意されている。NFC 認証は、カードの IC チップに記録された情報を NFC 部 23A で読み取ることによって実行される所持認証である。画像認証は、カードを撮影部 26 で撮影することによって実行される所持認証である。以降、NFC 認証と画像認証を区別しないときは、単に所持認証と記載する。

40

【0024】

図 3 では、NFC 認証の流れが示されている。ユーザが開始画面 G 5 のボタン B 50 を選択すると、NFC 部 23A が起動し、カードの IC チップに記録された情報を NFC 部 23A で読み取るための読取画面 G 6 が表示部 25 に表示される。なお、利用登録時に所持認証が実行されてもよく、この場合には、利用登録時に読取画面 G 6 が表示されてもよい。読取画面 G 6 が表示されると、ユーザは、自身が所持するカードにユーザ端末 20 を近づける。

50

【0025】

図4は、カードのICチップをNFC部23Aで読み取る様子の一例を示す図である。図4のカードC1は、第1実施形態の説明のために用意した架空のものである。図4に示すように、ユーザがカードC1のICチップcpにユーザ端末20を近づけると、NFC部23Aは、ICチップcpに記録された情報を読み取る。NFC部23Aは、ICチップcp内の任意の情報を読み取り可能である。第1実施形態では、NFC部23Aが、ICチップcpに記録された個人番号を読み取る場合を説明する。

【0026】

ユーザ端末20は、サーバ10に、ICチップcpから読み取った個人番号を送信する。この個人番号は、ユーザ端末20からサーバ10に入力されるので、以降では、この個人番号を入力個人番号と記載する。第1実施形態における入力とは、サーバ10に何らかのデータを送信することを意味する。サーバ10には、正解となる個人番号が予め利用登録時に登録されている。以降、この個人番号を登録個人番号と記載する。なお、入力個人番号と登録個人番号を特に区別しないときは、単に個人番号と記載することがある。

【0027】

サーバ10は、ユーザ端末20から入力個人番号を受信する。ユーザがカードC1の正当な持ち主であれば、入力個人番号と、ログイン中のユーザの登録個人番号と、が一致する。入力個人番号と、ログイン中のユーザの登録個人番号と、が一致する場合、図3に示すように、所持認証が成功したことを示す成功画面G7が表示部25に表示される。成功画面G7に示すように、所持認証を成功させたユーザ端末20から利用可能なサービスが増える。

【0028】

一方、入力個人番号と、ログイン中のユーザの登録個人番号と、が一致しない場合、所持認証が失敗したことを示す失敗画面G8が表示部25に表示される。この場合、ユーザ端末20から利用可能なサービスは制限されたままとなる。ユーザは、読取画面G6に戻ってカードC1の読み取りを再度実行したり、コールセンターに問い合わせたりする。第三者が不正にログインしていれば、手元にカードC1がなく、所持認証を成功させることができないので、第三者のユーザ端末20から利用可能なサービスは制限される。

【0029】

画像認証も同様の流れで実行される。NFC認証ではNFC部23Aを利用して入力個人番号が取得されるのに対し、画像認証では、カードC1が撮影された撮影画像を利用して入力個人番号が取得される。例えば、ユーザが開始画面G5のボタンB51を選択すると、撮影部26が起動する。撮影部26は、カードC1を撮影する。ユーザ端末20は、サーバ10に、撮影画像を送信する。サーバ10は、撮影画像を受信すると、撮影画像に光学文字認識を実行して入力個人番号を取得する。入力個人番号が取得された後の流れは、NFC認証と同様である。

【0030】

なお、光学文字認識は、ユーザ端末20で実行されてもよい。また、撮影画像から入力個人番号を取得する方法は、光学文字認識に限られない。この方法自体は、公知の種々の方法を利用可能である。例えば、入力個人番号を含むコード（例えば、バーコード又は二次元コード）がカードC1に形成されている場合、撮影画像に撮影されたコードを利用して入力個人番号が取得されてもよい。コードから入力個人番号を取得する処理は、サーバ10により実行されてもよいし、ユーザ端末20により実行されてもよい。

【0031】

以上のように、第1実施形態では、所持認証を成功させたユーザ端末20から利用可能なサービスは、所持認証を成功させていないユーザ端末20から利用可能なサービスよりも多くなる。第三者は、ユーザID及びパスワードを不正に入手して不正にログインしても、カードC1を所持しておらず所持認証を成功させることができないので、利用可能なサービスは制限される。このため、第三者によるサービスの不正利用を抑制し、サービスにおけるセキュリティが高まる。

【0032】

ただし、第三者によるサービスの不正利用を抑止したとしても、第三者は、少ない種類の中でサービスを不正に利用することがある。例えば、図2の例であれば、第三者は、他人になりすまして、証明書を請求したり窓口を予約したりすることがある。そこで、第1実施形態では、サービスにおける不正を検知するための学習モデルを利用して、第三者による不正を検知するようにしている。

【0033】

学習モデルは、機械学習を利用したモデルである。機械学習は、人工知能と呼ばれることもある。機械学習自体は、公知の種々の方法を利用可能であり、例えば、ニューラルネットワークを利用可能である。広い意味では、深層学習又は強化学習も機械学習に分類されるので、学習モデルは、深層学習又は強化学習を利用して作成されたモデルであってもよい。学習モデルは、機械学習を利用したルール又は決定木のモデルであってもよい。本実施形態では、教師有り学習を例に挙げるが、教師無し学習又は半教師有り学習であってもよい。

【0034】

第1実施形態の学習モデルは、他人のユーザIDで不正にログインした第三者の不正だけでなく、自身のユーザIDでログインしたユーザの不正も検知可能である。例えば、ユーザが自身のユーザIDでサービスにログインし、いたずら目的で証明書を大量に請求したり、窓口を予約して無断キャンセルをしたりすることもある。このような不正をするユーザの行動に一定の傾向があれば、学習モデルは、この傾向を学習することによって、不正を検知可能である。

【0035】

図5は、学習モデルの一例を示す図である。図5に示すように、第1実施形態では、教師有り学習を利用した学習モデルMを例に挙げる。教師有り学習では、学習モデルMへの入力と、学習モデルMから取得したい理想的な出力と、の関係が定義された訓練データが、学習モデルMに学習される。第1実施形態の学習モデルMは、不正であることを示す第1の値、又は、正当であることを示す第2の値を出力する場合を説明するが、不正の疑いを示すスコアを出力してもよい。スコアが出力される場合は、後述の変形例で説明する。第1実施形態の学習モデルMは、不正であるか否かを分類することになる。即ち、学習モデルMは、不正であるか否かのラベリングを行う。

【0036】

訓練データは、学習モデルMの作成者が手動で作成することが多い。学習モデルMの精度を高めようとすると、多数の訓練データを準備する必要がある。その全てを管理者が手動で作成するのは、非常に手間がかかる。例えば、管理者は、サービスにおける個々の行動が正当であるか不正であるかを判断し、訓練データを作成しなければならない。

【0037】

この点、所持認証を実行したユーザは、所持認証の実行に必要な物理的なカードを所有しているため、不正をしない確率が非常に高い。不正をするユーザは、フィッシング等により、ユーザID及びパスワードを不正に入手することはできても、物理的なカードを盗むことはできずに、所持認証を実行せずにサービスを利用する確率が非常に高い。仮に、不正をするユーザが物理的なカードを盗めたとしても、所持認証を実行したユーザが不正をすると、誰が不正をしたかを簡単に特定できるので、不正をするユーザは、身元を隠すために、所持認証を実行せずにサービスを利用する確率が非常に高い。例えば、不正をするユーザは、自身の個人番号ではない番号を個人番号として入力して利用登録を完了させることがある。図2及び図3のような流れでサービスが提供される場合、自分の個人番号ではない番号が入力されたとしても、制限された中でサービスが利用可能になる。

【0038】

そこで、第1実施形態では、所持認証を実行したユーザの行動を正当とみなし、訓練データを作成するようにしている。以降、所持認証を実行したユーザを、認証済みユーザと記載する。図5に示すように、第1実施形態の訓練データは、認証済みユーザの行動に基

10

20

30

40

50

づいて作成されている。図5の例では、訓練データは、場所情報、日時情報、及び利用情報を含む入力部分と、正当であることを示す出力部分と、を含む。

【0039】

場所情報は、ユーザ端末20の場所を示す。場所は、任意の情報によって示されてよく、例えば、緯度経度、住所、携帯基地局情報、無線LANのアクセスポイント情報、又はIPアドレスによって示される。場所情報は、サービスを普段利用する中心地からの距離であってもよい。中心地は、あるユーザIDから利用された場所の平均値であってもよいし、あるユーザ端末20から利用された場所の平均値であってもよい。日時情報は、サービスの利用日時を示す。利用情報は、サービスがどのように利用されたかを示す。利用情報は、サービスの利用履歴ということもできる。例えば、利用情報は、利用されたサービスの種類、利用内容、ユーザの操作、又はこれらの組み合わせを示す。

10

【0040】

例えば、サーバ10は、学習済みの学習モデルMを利用して、サービスにログイン中のユーザの不正を検知する。以降、不正検知の対象となるユーザを、対象ユーザと記載する。学習モデルMには、対象ユーザの場所情報、日時情報、及び利用情報を含む対象情報が入力される。学習モデルMは、対象情報に基づいて、不正であるか否かの推定結果を出力する。学習モデルMからの出力が不正を示していれば、対象ユーザに対するサービスの提供が制限される。学習モデルMからの出力が正当を示していれば、対象ユーザに対するサービスの提供は制限されない。

【0041】

以上のように、第1実施形態の不正検知システムSは、不正をしない確率が非常に高い認証済みユーザの認証済み情報に基づいて、教師有り学習を利用した学習モデルMに学習させる訓練データを作成する。これにより、学習モデルMの作成者が手動で訓練データを作成するといった手間を省き、学習モデルMの作成を簡易化するようにしている。以降、第1実施形態の詳細を説明する。

20

【0042】

[1-3. 第1実施形態において実現される機能]

図6は、第1実施形態の不正検知システムSで実現される機能の一例を示す機能ブロック図である。ここでは、サーバ10及びユーザ端末20の各々で実現される機能を説明する。

30

【0043】

[1-3-1. サーバにおいて実現される機能]

図6に示すように、サーバ10では、データ記憶部100、認証済み情報取得部101、作成部102、及び不正検知部103が実現される。データ記憶部100は、記憶部12を主として実現される。認証済み情報取得部101、作成部102、及び不正検知部103の各々は、制御部11を主として実現される。

【0044】

[データ記憶部]

データ記憶部100は、学習モデルMを作成するために必要なデータを記憶する。例えば、データ記憶部100は、ユーザデータベースDB1、訓練データベースDB2、及び学習モデルMを記憶する。

40

【0045】

図7は、ユーザデータベースDB1のデータ格納例を示す図である。図7に示すように、ユーザデータベースDB1は、利用登録が完了したユーザに関する情報が格納されたデータベースである。例えば、ユーザデータベースDB1には、ユーザID、パスワード、氏名、住所、電話番号、登録個人番号、端末ID、所持認証フラグ、サービスの利用設定、場所情報、日時情報、及び利用情報が格納される。

【0046】

例えば、ユーザが利用登録をすると、ユーザデータベースDB1に新たなレコードが作成される。このレコードには、利用登録時に指定されたユーザID、パスワード、氏名、

50

住所、電話番号、及び登録個人番号が格納される。第1実施形態では、登録個人番号は、利用登録後に変更できないものとする。このため、第三者は、不正にログインしたとしても、勝手に登録個人番号を変更できないものとする。利用登録時には、個人番号の確認は行われないため、不正をするユーザは、自身の個人番号ではない番号を個人番号として入力して利用登録を完了させることがある。

【0047】

端末IDは、ユーザ端末20を識別可能な情報である。第1実施形態では、サーバ10が端末IDを発行する場合を説明する。端末IDは、所定のルールに基づいて発行される。サーバ10は、他の端末IDと重複しないように、端末IDを発行する。端末IDは、有効期限が設定されてもよい。端末IDは、任意のタイミングで発行可能である。例えば、アプリが起動したタイミング、端末IDに設定された有効期限が切れたタイミング、又は端末IDを更新するための操作が行われたタイミングで端末IDが発行される。

10

【0048】

なお、ユーザ端末20は、端末ID以外の任意の情報によって識別可能である。例えば、端末ID以外にも、IPアドレス、Cookieに格納された情報、SIMカードに格納されたID、ICチップ27に格納されたID、又はユーザ端末20の個体識別情報によってユーザ端末20が識別されてもよい。ユーザ端末20を何らかの識別可能な情報がユーザデータベースDB1に格納されるようにすればよい。

【0049】

ユーザIDに関連付けられた端末IDは、このユーザIDからログインしたことのあるユーザ端末20の端末IDである。このため、あるユーザIDの正当な持ち主であるユーザが新たなユーザ端末20からログインすれば、このユーザ端末20の端末IDが、このユーザIDに関連付けられる。第三者が、このユーザIDから不正にログインした場合も、第三者のユーザ端末20の端末IDが、このユーザIDに関連付けられる。

20

【0050】

端末IDには、所持認証フラグ、利用設定、時間情報、場所情報、日時情報、及び利用情報が関連付けられる。第1実施形態では、ユーザID及び端末IDの組み合わせごとに、所持認証フラグ等の情報が関連付けられる。図7の例であれば、ユーザID「taro.yamada123」は、2台のユーザ端末20からログインされたことがある。ユーザID「hanako.suzuki999」は、3台のユーザ端末20からログインされたことがある。ユーザID「kimura9876」は、1台のユーザ端末20からのみログインされたことがある。

30

【0051】

所持認証フラグは、所持認証が実行されたか否かを示す情報である。例えば、所持認証フラグが「1」であることは、NFC認証が実行されたことを示す。所持認証フラグが「2」であることは、画像認証が実行されたことを示す。所持認証フラグが「0」であることは、所持認証が実行されていないことを示す。第1実施形態では、利用登録時に所持認証が実行されない場合を説明するので、所持認証フラグの初期値は「0」になる。利用登録後に所持認証が実行されると、所持認証フラグが「1」又は「2」に変わる。利用登録時に所持認証を実行可能とする場合には、ユーザが利用登録時に所持認証を実行すれば、所持認証フラグの初期値は「1」又は「2」になる。

40

【0052】

利用設定では、アプリから利用可能なサービスの種類が示される。所持認証フラグ「1」又は「2」の利用設定は、所持認証フラグ「0」の利用設定よりも、利用可能なサービスが多くなる。所持認証の実行有無及び利用設定の関係（即ち、所持認証フラグ及び利用設定の関係）は、データ記憶部100に予め定義されているものとする。図6の例であれば、所持認証フラグ「1」又は「2」の利用設定は、全てのサービスを利用可能な設定になる。所持認証フラグ「0」の利用設定は、一部のサービスのみ利用可能な設定になる。

【0053】

場所情報、日時情報、及び利用情報の詳細は、先述した通りである。あるユーザ端末2

50

0 からあるユーザIDでログインした状態でサービスが利用されると、このユーザID及びこのユーザ端末20の組み合わせに関連付けられた場所情報、日時情報、及び利用情報が更新される。場所情報を取得する方法自体は、GPSや携帯基地局等を利用した公知の方法を利用可能である。日時情報を取得する方法自体も、リアルタイムクロック等を利用した公知の方法を利用可能である。利用情報は、サービスに応じた情報が格納されるようにすればよく、詳細な内容は先述した通りである。

【0054】

図8は、訓練データベースDB2のデータ格納例を示す図である。図8に示すように、訓練データベースDB2は、学習モデルMに学習させる訓練データが格納されたデータベースである。本実施形態では、学習モデルMに対する入力部分と、正解となる出力部分と、のペアを訓練データ（教師データ）と記載する。図8の出力部分の例では、正当を「0」で示す。不正は、他の値であればよく、例えば、「1」である。訓練データベースDB2には、このペアの集まりが格納される。訓練データの詳細は、図5で説明した通りである。訓練データは、作成部102により作成される。一部の訓練データは、学習モデルMの作成者が手動で作成してもよいし、公知の訓練データの作成ツールを利用して作成されてもよい。

10

【0055】

データ記憶部100は、学習済みの学習モデルMのプログラム及びパラメータを記憶する。データ記憶部100は、訓練データが学習される前の学習モデルMと、訓練データの学習に必要なプログラムと、を記憶してもよい。データ記憶部100が記憶するデータは、上記の例に限られない。データ記憶部100は、任意のデータを記憶可能である。

20

【0056】

学習モデルMは、機械学習を利用したモデルである。機械学習は、人工知能と呼ばれることもある。機械学習自体は、公知の種々の方法を利用可能であり、例えば、ニューラルネットワークを利用可能である。広い意味では、深層学習又は強化学習も機械学習に分類されるので、学習モデルMは、深層学習又は強化学習を利用して作成されたモデルであってもよい。本実施形態では、教師有り学習を例に挙げるが、教師無し学習又は半教師有り学習であってもよい。

【0057】

〔認証済み情報取得部〕

認証済み情報取得部101は、所定のサービスを利用可能なユーザ端末20から所定の認証を実行した認証済みユーザの行動に関する認証済み情報を取得する。第1実施形態では、この認証が、ユーザ端末20を利用して、所定のカードC1を所持しているか否かを確認するための所持認証である場合を例に挙げる。このため、所持認証について説明している箇所は、所定の認証と読み替えることができる。即ち、NFC認証又は画像認証について説明している箇所は、所定の認証と読み替えることができる。第1実施形態では、認証済みユーザがユーザ端末20から所持認証を実行したユーザである場合を説明するが、認証済みユーザは、ユーザ端末20から所定の認証を実行したユーザであればよい。

30

【0058】

所定の認証は、ユーザ端末20から実行可能な認証である。所定の認証は、ログイン時の認証であってもよいが、第1実施形態では、所定の認証は、ログイン時の認証とは異なる認証であるものとする。所定の認証は、カードC1を利用した所持認証に限られない。所定の認証は、種々の認証方法を利用可能である。例えば、所定の認証は、カードC1以外の所持物を確認する所持認証であってもよい。この所持物は、本人確認が可能な任意のものであればよい。例えば、所持物は、パスポートのようなカード以外の身分証明書、何らかの認証情報が記録された情報記憶媒体、又は何らかの認証情報が形成された紙であってもよい。例えば、所持物は、認証情報を含むコードのような電子的な物であってもよい。

40

【0059】

所定の認証は、所持認証に限られない。例えば、所定の認証は、パスワード認証、パスコード認証、暗証番号認証、又は合言葉認証といった知識認証であってもよい。所定の認

50

証がパスワード認証の場合には、ログイン時とは異なるパスワードが利用されるものとする。例えば、所定の認証は、顔認証、指紋認証、又は虹彩認証といった生体認証であってもよい。第1実施形態では、所定の認証がログイン時の認証よりもセキュアなものである場合を説明するが、ログイン時の認証の方が所定の認証よりもセキュアであってもよい。ログイン時の認証も、パスワード認証に限られず、任意の認証方法であってもよい。

【0060】

第1実施形態の所持認証で利用されるカードC1は、所持認証で利用される入力個人番号を含む。例えば、入力個人番号は、カードC1のICチップcpに電子的に記録されている。第1実施形態では、入力個人番号は、カードC1の表面にも形成されている。所持認証で正解となる登録個人番号は、ユーザデータベースDB1に登録されている。入力個人番号及び登録個人番号の各々は、認証時に利用される認証情報の一例である。

10

【0061】

なお、所定の認証として他の認証方法が利用される場合には、認証方法に応じた認証情報が用いられるようにすればよい。例えば、知識認証が利用されるのであれば、認証情報は、パスワード、パスコード、暗証番号、又は合言葉であってもよい。生体認証が利用されるのであれば、認証情報の各々は、顔写真、顔の特徴量、指紋パターン、又は虹彩パターンであってもよい。

【0062】

例えば、NFC認証を利用して所持認証が実行される場合、サーバ10は、ユーザ端末20から、NFC部23Aを利用して取得されたカードC1の入力個人番号を取得する。サーバ10は、ユーザデータベースDB1を参照し、ユーザ端末20から取得した入力個人番号と、ログイン中のユーザIDに関連付けられた登録個人番号と、が一致するか否かを判定する。これらが一致する場合、所持認証は成功する。これらが一致しない場合、所持認証は失敗する。

20

【0063】

例えば、画像認証を利用して所持認証が実行される場合、サーバ10は、ユーザ端末20から、カードC1が撮影された撮影画像を取得する。サーバ10は、光学文字認識を利用し、撮影画像から入力個人番号を取得する。入力個人番号が取得された後の所持認証の流れは、NFC認証と同様である。第1実施形態では、入力個人番号がカードC1の表面に印刷されている場合を説明するが、入力個人番号は、カードC1の表面にエンボス加工された凹凸として形成されていてもよい。入力個人番号は、カードC1の表面及び裏面の少なくとも一方に形成されていればよい。

30

【0064】

第1実施形態のサービスは、複数のユーザ端末20の各々から、同じユーザIDでログイン可能である。認証部101は、ユーザ端末20ごとに、当該ユーザ端末20からユーザIDでサービスにログインした状態で所持認証を実行可能である。例えば、図7のユーザID「taro.yamada123」のユーザが、2台のユーザ端末20を利用していたとする。これら2台のユーザ端末20を、第1ユーザ端末20A及び第2ユーザ端末20Bと記載する。

【0065】

サーバ10は、第1ユーザ端末20Aから、ユーザID「taro.yamada123」でサービスにログインした状態で所持認証を実行可能である。認証部101は、第2ユーザ端末20Bから、同じユーザID「taro.yamada123」でサービスにログインした状態で所持認証を実行可能である。1人のユーザが3台以上のユーザ端末20を利用する場合も同様に、認証部101は、個々のユーザ端末20ごとに、所持認証を実行可能である。先述したように、所持認証を実行するか否かは、ユーザの任意なので、全てのユーザ端末20で所持認証を実行しなければならないわけではない。

40

【0066】

認証済み情報は、認証済みユーザの行動に関する情報である。行動とは、ユーザ端末20に対する操作内容、ユーザ端末20からサーバ10に送信された情報、又はこれらの組

50

み合わせである。別の言い方をすれば、行動は、どのようにサービスを利用したかを示す情報である。第1実施形態では、場所情報、日時情報、及び利用情報の組み合わせが行動に関する情報に相当する。認証済みユーザの場所情報、日時情報、及び利用情報の組み合わせは、認証済み情報の一例である。このため、以降では、この組み合わせを認証済み情報と記載する。

【0067】

なお、認証済み情報は、第1実施形態の例に限られず、認証済みユーザの何らかの行動に関する情報であればよい。即ち、認証済み情報は、不正であるか否かと何らかの相関関係のある特徴であればよい。例えば、認証済み情報は、ユーザがログインしてから所定の画面に到達するまでの時間、この画面に到達するまでに表示された画面の数若しくは種類、ある1つの画面に対する操作の数、ポインタの軌跡、又はこれらの組み合わせであってもよい。認証済み情報は、サービスに応じた情報であればよい。認証済み情報の他の例は、後述の変形例で説明する。

【0068】

第1実施形態では、認証済み情報は、ユーザデータベースDB1に格納されている。図7の例であれば、所持認証フラグが「1」又は「2」のレコードに格納された場所情報、日時情報、及び利用情報の組み合わせが認証済み情報に相当する。認証済み情報取得部101は、ユーザデータベースDB1を参照し、認証済み情報を取得する。第1実施形態では、認証済み情報取得部101は、複数の認証済み情報を取得する場合を説明するが、認証済み情報取得部101は、少なくとも1つの認証済み情報を取得すればよい。

【0069】

第1実施形態では、認証済み情報取得部101は、日時情報が示す日時が直近の所定期間（例えば、1週間～1月程度）の認証済み情報を取得する場合を説明するが、ユーザデータベースDB1に格納された全ての認証済み情報を取得してもよい。認証済み情報取得部101は、所定期間内の全ての認証済み情報を取得しなくてもよく、所定期間内の一部の認証済み情報をランダムに選択して取得してもよい。認証済み情報取得部101は、学習モデルMの学習に十分な数の認証済み情報を取得すればよい。

【0070】

【作成部】

作成部102は、認証済み情報に基づいて、認証済みユーザの行動が正当と推定されるように、サービスにおける不正を検知するための学習モデルMを作成する。学習モデルMを作成するとは、学習モデルMの学習を行うことである。学習モデルMのパラメータを調整することは、学習モデルMを作成することに相当する。パラメータ自体は、公知の機械学習で利用されるものであればよく、例えば、重み係数やバイアス等である。学習モデルMの学習方法自体は、種々の手法を利用可能であり、例えば、深層学習又は強化学習の手法を利用可能である。他にも例えば、勾配降下法が利用されてもよいし、深層学習であれば誤差逆伝播法が利用されてもよい。

【0071】

第1実施形態では、学習モデルMは、教師有り学習のモデルである。作成部102は、認証済み情報に基づいて、認証済みユーザの行動が正当であることを示す訓練データを作成する。この訓練データは、第1訓練データの一例である。後述の変形例の説明では、他の訓練データを説明するので、第1訓練データや第2訓練データといったように個々の訓練データを区別するが、第1実施形態では、他の訓練データの説明はしないので、第1訓練データを単に訓練データと記載する。

【0072】

例えば、作成部102は、認証済み情報である入力部分と、正当を示す出力部分と、を含む訓練データを作成する。入力部分は、任意の形式で表現可能であり、例えば、ベクトル形式、配列形式、又は単一の数値で表現されてよい。認証済み情報に含まれる場所情報、日時情報、及び利用情報に含まれる項目が数値化されたものが入力部分であるものとする。この数値化は、学習モデルMの内部で行われてもよい。入力部分は、行動の特徴量に

相当する。出力部分は、学習モデルMの出力の正解に相当する。

【0073】

作成部102は、認証済み情報ごとに訓練データを作成し、訓練データベースDB2に格納する。作成部102は、訓練データに基づいて、学習モデルMを学習させることによって、学習モデルMを作成する。作成部102は、訓練データの入力部分が入力された場合に、訓練データの出力部分が取得されるように、学習モデルMを学習させる。作成部102は、訓練データベースDB2に格納された全ての訓練データを利用して学習モデルMを作成してもよいし、一部の訓練データのみを利用して学習モデルMを作成してもよい。

【0074】

〔不正検知部〕

不正検知部103は、作成済みの学習モデルMを利用して、不正検知を行う。不正検知部103は、対象ユーザがサービスにログインすると、対象ユーザの場所情報、日時情報、及び利用情報を取得してユーザデータベースDB1に格納する。これらの情報の組み合わせは、図5に示す対象情報である。不正検知部103は、所定の不正検知のタイミングが訪れると、対象ユーザの対象情報に基づいて、学習モデルMの出力を取得する。第1実施形態では、不正検知部103は、学習モデルMに対象情報を入力し、学習モデルMからの出力を取得する場合を説明するが、不正検知部103は、対象情報に何らかの演算や数値化の処理を実行したうえで、当該処理が実行された対象情報を学習モデルMに入力してもよい。

【0075】

不正検知部103は、学習モデルMの出力が不正を示していれば、対象ユーザに対するサービスの提供、即ち対象ユーザによるサービスの利用を制限する。不正検知部103は、この出力が正当を示していれば、対象ユーザによるサービスの利用を制限しない。不正検知のタイミングは、任意のタイミングであってよく、例えば、トップ画面G3のボタンB30が選択された場合、ユーザデータベースDB1に登録された情報が変更される場合、サービスへのログイン時、又は何らの決済処理が実行される場合であってもよい。

【0076】

〔1-3-2. ユーザ端末において実現される機能〕

図5に示すように、ユーザ端末20では、データ記憶部200、表示制御部201、及び受付部202が実現される。データ記憶部200は、記憶部22を主として実現される。表示制御部201及び受付部202の各々は、制御部21を主として実現される。データ記憶部200は、第1実施形態で説明する処理に必要なデータを記憶する。例えば、データ記憶部200は、アプリを記憶する。表示制御部201は、アプリに基づいて、図2及び図3で説明した各画面を表示部25に表示させる。受付部202は、各画面に対するユーザの操作を受け付ける。ユーザ端末20は、サーバ10に、ユーザの操作内容を送信する。他にも例えば、ユーザ端末20は、認証済み情報の取得に必要な場所情報等を送信する。

【0077】

〔1-4. 第1実施形態において実行される処理〕

図9は、第1実施形態において実行される処理の一例を示すフロー図である。図9に示す処理は、制御部11、21が、それぞれ記憶部12、22に記憶されたプログラムに従って動作することによって実行される。この処理は、図6に示す機能ブロックにより実行される処理の一例である。この処理が実行されるにあたり、ユーザの利用登録が完了しているものとする。ユーザ端末20は、サーバ10により発行された端末IDを予め記憶しているものとする。

【0078】

図9に示すように、サーバ10は、ユーザデータベースDB1に基づいて、認証済みユーザの認証済み情報を取得する(S100)。S100では、サーバ10は、所持認証フラグが「1」又は「2」のレコードのうち、日時情報が示す日時が直近の所定期間であるレコードに格納された認証済み情報を取得する。

【0079】

サーバ10は、S100で取得した認証済み情報に基づいて、訓練データを作成する（S101）。S101では、サーバ10は、認証済み情報である入力部分と、不正を示す出力部分と、を含む訓練データを作成し、訓練データベースDB2に格納する。サーバ10は、訓練データの作成が完了したか否かを判定する（S102）。S102では、サーバ10は、所定数の訓練データを作成したか否かを判定する。

【0080】

訓練データの作成が完了したと判定されない場合（S102；N）、S100の処理に戻り、新たに訓練データが作成されて訓練データベースDB2に格納される。S102において、訓練データの作成が完了したと判定された場合（S102；Y）、サーバ10は、訓練データベースDB2に基づいて、学習モデルMを作成する（S103）。S103では、サーバ10は、訓練データベースDB2に格納された個々の訓練データの入力部分が入力された場合に、この訓練データの出力部分が出力されるように、個々の訓練データを学習モデルMに学習させる。

10

【0081】

S103で学習モデルMが作成されると、サービスにおける不正検知で利用可能になる。ユーザ端末20は、対象ユーザの操作に基づいてアプリを起動させ、トップ画面G3を表示部25に表示させる（S104）。アプリの起動時には、サーバ10及びユーザ端末20の間でログインが実行されてもよい。ログインでは、ユーザID及びパスワードの入力が要求されてもよいし、過去にログイン済みであることを示す情報をユーザ端末20に記憶させておき、この情報がログインで利用されてもよい。以降、ユーザ端末20が何らかの形でサーバ10にアクセスすると、ユーザ端末20の端末IDに関連付けられた場所情報、日時情報、及び利用情報が適宜更新される。なお、サーバ10は、ログインが成功してトップ画面G3が表示される前に、ユーザ端末20の端末IDに関連付けられた利用設定に基づいて、利用できないサービスのボタンB30を選択できないようなトップ画面G3の表示データを生成し、ユーザ端末20に送信してもよい。

20

【0082】

ユーザ端末20は、操作部24の検出信号に基づいて、対象ユーザの操作を特定する（S105）。S105では、行政サービスを利用するためのボタンB30の選択、又は、所持認証を実行するためのボタンB31の選択の何れかが行われる。所持認証を実行済みのユーザ端末20であれば、ボタンB31を選択できないようにしてもよい。なお、対象ユーザがアプリを終了するための操作やアプリをバックグラウンドに移行させるための操作を行った場合（S105；終了）、本処理は終了する。

30

【0083】

S105において、ボタンB30が選択された場合（S105；B30）、ユーザ端末20は、サーバ10に、対象ユーザがボタンB30から選択した種類のサービスの提供を要求する（S106）。サーバ10は、対象ユーザの対象情報を学習モデルMに入力し、学習モデルMからの出力を取得する（S107）。なお、ここでは、対象ユーザがログインした後にS107の処理が実行される場合を説明するが、対象ユーザがログインする時にS107の処理が実行されてもよい。この場合、不正なログインを検知し、不正なログインが発生することを防止できる。対象情報は、対象ユーザ（即ち、ログイン中のユーザ）の場所情報、日時情報、及び利用情報である。対象ユーザが複数のユーザ端末20からログインしたことがあれば、ログイン中のユーザ端末20の端末IDに関連付けられた対象情報に基づいて、学習モデルMからの出力が取得される。

40

【0084】

サーバ10は、学習モデルMからの出力を参照する（S108）。学習モデルMからの出力が不正を示す場合（S108；不正）、サーバ10は、サービスの提供を制限する（S109）。S109では、サーバ10は、ユーザが選択した種類のサービスを提供しない。ユーザ端末20には、エラーメッセージが表示される。学習モデルMからの出力が正当を示す場合（S108；正当）、サーバ10と、ユーザ端末20と、の間でサービスを

50

提供するためのサービス提供処理が実行され（S 1 1 0）、本処理は終了する。S 1 1 0では、サーバ10は、ユーザデータベースDB1を参照し、ログイン中のユーザのユーザIDと、ユーザ端末20の端末IDと、に関連付けられた利用設定を取得する。サーバ10は、この利用設定に基づいて、サービスを提供する。サーバ10は、ユーザ端末20からユーザの操作内容を受信し、操作内容に応じた処理を実行する。

【0085】

S 1 0 5において、ボタンB 3 1が選択された場合（S 1 0 8；B 3 1）、ユーザ端末20は、開始画面G 5を表示部25に表示させ、サーバ10及びユーザ端末20の間で所持認証が実行され（S 1 1 1）、本処理は終了する。S 1 1 1においてN F C認証が選択された場合には、ユーザ端末20は、N F C部23Aにより読み取られた入力個人番号を、サーバ10に送信する。サーバ10は、入力個人番号を受信すると、ユーザデータベースDB1を参照し、受信した入力個人番号と、ログイン中のユーザの登録個人番号と、が一致するか否かを判定する。サーバ10は、これらが一致した場合には、所持認証が成功したと判定し、所持認証フラグを「1」にしてサービスの利用制限が解除されるように、利用設定を変更する。画像認証が選択された場合には、撮影画像から入力個人番号が取得され、N F C認証と同様の流れで画像認証が実行される。この場合の所持認証フラグは「2」になる。

【0086】

第1実施形態の不正検知システムSによれば、認証済み情報に基づいて、認証済みユーザの行動が正当と推定されるように、学習モデルMを作成する。認証済みユーザが正当である確率が非常に高いことに着目することによって、学習モデルMの作成者が手動で訓練データを作成しなくても学習モデルMを作成できるので、学習モデルMの作成を簡易化できる。また、訓練データの作成から学習モデルMの学習までの一連の処理を自動化し、学習モデルMを迅速に作成できる。不正検知システムSに、最新の傾向を学習させた学習モデルMを迅速に適用し、精度良く不正を検知できる。その結果、サービスにおける不正利用を防止し、セキュリティが高まる。本来は正当なはずの対象ユーザの行動が不正と推定されてしまいサービスを利用できなくなる、といった利便性の低下も防止できる。学習モデルMに認証済みユーザの正当な行動だけを学習させたとしても、不正な行動は、正当な行動の特徴とは異なることが多いので、学習モデルMは、正当な行動とは特徴が異なる行動を検知することによって、不正を検知できる。

【0087】

また、不正検知システムSは、所持認証を実行した認証済みユーザの認証済み情報を利用して学習モデルMを作成することによって、正当である確率が非常に高い認証済みユーザの認証済み情報を利用して、精度の高い学習モデルMを作成できる。精度の高い学習モデルMを作成することによって、サービスにおける不正利用をより確実に防止し、セキュリティが効果的に高まる。本来は正当なはずの対象ユーザの行動が不正と推定されてしまいサービスを利用できなくなる、といったこともより確実に防止できる。

【0088】

また、不正検知システムSは、認証済み情報に基づいて、認証済みユーザの行動が正当であることを示す訓練データを作成し、この訓練データに基づいて、学習モデルMを学習させることによって、訓練データを自動的に作成し、学習モデルMの作成者の手間を軽減できる。学習モデルMの作成で最も手間がかかる工程の1つである訓練データの作成を自動化することによって、学習モデルMを迅速に作成できる。その結果、サービスにおける不正利用をより確実に防止し、セキュリティが効果的に高まる。

【0089】

[2. 第2実施形態]

次に、本開示に係る学習モデルM評価システムの実施形態の一例である第2実施形態を説明する。第2実施形態では、学習モデルM評価システムを、第1実施形態で説明した不正検知システムSに適用した場合を例に挙げる。このため、第2実施形態で不正検知システムSと記載した箇所は、学習モデルM評価システムと読み替えることができる。学習モ

デルM評価システムは、学習モデルMの評価までを行い、不正検知は、他のシステムで実行されてもよい。即ち、学習モデルM評価システムは、不正検知システムSのうちの不正検知の機能を含まなくてもよい。

【0090】

また、第1実施形態と同様にして学習モデルMが作成される場合を説明するが、第2実施形態の学習モデルMは、第1実施形態とは異なる方法で作成されてもよい。例えば、学習モデルMの作成者が手動で作成した訓練データに基づいて、学習モデルMが作成されてもよい。他にも例えば、公知の訓練データの作成支援ツールを利用して作成された訓練データに基づいて、学習モデルMが作成されてもよい。このため、第2実施形態の不正検知システムSは、第1実施形態で説明した機能を含まなくてもよい。なお、第2実施形態では、第1実施形態と同様の点については説明を省略する。

10

【0091】

[2-1. 第2実施形態の概要]

不正検知システムSにおけるユーザの行動は、日々変化するので、学習モデルMに最近の傾向が学習されていないと、学習モデルMの不正検知の精度が徐々に落ちることがある。この点は、第1実施形態以外の方法で作成された学習モデルMも同様である。教師無し学習又は半教師有り学習が利用される場合も同様である。そこで、第2実施形態では、認証済みユーザが正当である確率が非常に高いことに着目し、認証済み情報に基づいて、学習モデルMの精度を正確に評価するようにしている。

【0092】

20

図10は、第2実施形態の概要を示す図である。図10に示すように、複数の認証済み情報の各々が、学習モデルMに入力される。認証済み情報は、正当である確率が非常に高い認証済みユーザの行動に関する情報なので、学習モデルMからの出力が正当を示していれば、学習モデルMの精度は低下していないと予測される。一方、学習モデルMからの出力が不正を示していれば、学習モデルMは、最近の認証済みユーザの行動（即ち、正当な行動）に対応できておらず、精度が低下している可能性がある。この場合、学習モデルMの作成者に精度が低下していることが通知されたり、最新の認証済み情報に基づいて、学習モデルMが作り直されたりする。

【0093】

以上のように、第2実施形態の不正検知システムSは、認証済み情報に基づいて、学習モデルMからの出力を取得し、認証済み情報に対応する出力に基づいて、学習モデルMの精度を評価する。正当である確率が非常に高い認証済みユーザの認証済み情報を利用することによって、学習モデルMの精度を正確に評価できる。以降、第2実施形態の詳細を説明する。

30

【0094】

[2-2. 第2実施形態において実現される機能]

図11は、第2実施形態の不正検知システムSで実現される機能の一例を示す機能ブロック図である。ここでは、サーバ10及びユーザ端末20の各々で実現される機能を説明する。

【0095】

40

[2-2-1. サーバにおいて実現される機能]

図11に示すように、サーバ10は、データ記憶部100、認証済み情報取得部101、作成部102、不正検知部103、出力取得部104、及び評価部105を含む。出力取得部104及び評価部105の各々は、制御部11を主として実現される。

【0096】

[データ記憶部、認証済み情報取得部、作成部、及び不正検知部]

データ記憶部100は、第1実施形態と同様である。第1実施形態の認証済み情報取得部101は、学習モデルMを作成するための認証済み情報を取得したが、第2実施形態の認証済み情報取得部101は、学習モデルMを評価するための認証済み情報を取得する。認証済み情報の利用目的が異なるだけであり、認証済み情報自体は同じである。認証済み

50

情報取得部 101 の他の点については、第 1 実施形態と同様である。作成部 102 及び不正検知部 103 も第 1 実施形態と同様である。

【0097】

〔出力取得部〕

出力取得部 104 は、認証済み情報に基づいて、サービスにおける不正を検知するための学習モデル M からの出力を取得する。例えば、出力取得部 104 は、複数の認証済み情報の各々に対応する出力を取得する。学習モデル M に認証済み情報が入力されて、学習モデル M からの出力が取得される処理は、第 1 実施形態で説明した通りである。認証済み情報に何らかの演算又は数値化の処理が実行されたうえで、当該処理が実行された認証済み情報が学習モデル M に入力されても良い点も、第 1 実施形態と同様である。

10

【0098】

〔評価部〕

評価部 105 は、認証済み情報に対応する出力に基づいて、学習モデル M の精度を評価する。認証済み情報に対応する出力とは、認証済み情報に基づいて取得された学習モデル M からの出力である。学習モデル M の精度とは、学習モデル M からどの程度所望の結果を得られるかを示す指標である。例えば、正当な行動の対象情報が入力された場合に、学習モデル M から正当を示す出力を取得できる確率は、学習モデル M の精度に相当する。不正な行動の対象情報が入力された場合に、学習モデル M から不正を示す出力を取得できる確率は、学習モデル M の精度に相当する。学習モデル M の精度は、任意の指標によって計測可能であり、例えば、正解率、適合率、再現率、F 値、特異度、偽陽性率、Log Loss、又は AUC (Area Under the Curve) を利用可能である。

20

【0099】

第 2 実施形態では、評価部 105 は、認証済み情報に対応する学習モデル M からの出力が正当を示す場合には、学習モデル M からの出力が不正を示す場合よりも、学習モデル M の精度が高いと評価する。例えば、評価部 105 は、複数の認証済み情報の各々に対応する出力に基づいて、学習モデル M の精度を評価する。評価部 105 は、学習モデル M に入力した認証済み情報のうち、学習モデル M からの出力が正当を示す割合を、正解率として計算する。評価部 105 は、正解率が高いほど、学習モデル M の精度が高いと評価する。即ち、評価部 105 は、正解率が低いほど、学習モデル M の精度が低いと評価する。学習モデル M の精度は、正解率ではなく、先述した種々の指標を利用可能である。

30

【0100】

〔2-2-2. ユーザ端末において実現される機能〕

図 11 に示すように、ユーザ端末 20 の機能は、第 1 実施形態と同様である。

【0101】

〔2-3. 第 2 実施形態において実行される処理〕

図 12 は、第 2 実施形態において実行される処理の一例を示すフロー図である。図 12 に示す処理は、制御部 11 が記憶部 12 に記憶されたプログラムに従って動作することによって実行される。この処理は、図 12 に示す機能ブロックにより実行される処理の一例である。

【0102】

図 12 に示すように、サーバ 10 は、ユーザデータベース DB1 を参照し、n (n は自然数) 個の認証済み情報を取得する (S200)。S200 では、サーバ 10 は、所持認証フラグが「1」又は「2」のレコードのうち、日時情報が示す日時が直近の所定期間であるレコードに格納された n 個の認証済み情報を取得する。サーバ 10 は、日時情報が示す日時が直近の所定期間である全ての認証済み情報を取得してもよいし、予め定められた個数の認証済み情報を取得してもよい。

40

【0103】

サーバ 10 は、S200 で取得した n 個の認証済み情報の各々に基づいて、学習モデル M からの n 個の出力を取得する (S201)。S201 では、サーバ 10 は、n 個の認証済み情報の各々を学習モデル M に次々と入力し、個々の認証済み情報に対応する出力を取

50

得する。サーバ10は、S201で取得したn個の出力のうち、正当を示す出力の割合を、学習モデルMの正解率として計算する(S202)。

【0104】

サーバ10は、学習モデルMの正解率が閾値以上であるか否かを判定する(S203)。学習モデルMの正解率が閾値以上であると判定された場合(S203; Y)、サーバ10は、学習モデルMの作成者に、学習モデルMの精度が高いことを示す評価結果を通知し(S204)、本処理は終了する。評価結果の通知は、任意の方法によって行われるようにすればよく、例えば、電子メール又は作成者が使用する管理プログラム内の通知を利用すればよい。S204の評価結果が通知された場合には、学習モデルMの精度が高いため、学習モデルMの作成者は、学習モデルMを作り直さない。この場合、現状の学習モデルMで不正検知が実行される。

10

【0105】

S203において、学習モデルMの正解率が閾値未満であると判定された場合(S203; N)、サーバ10は、学習モデルMの作成者に、学習モデルMの精度が低いことを示す評価結果を通知し(S205)、本処理は終了する。この場合、学習モデルMの精度が低いので、学習モデルMの作成者は、学習モデルMを作り直す。学習モデルMは、第1実施形態と同様の方法で作り返されてもよいし、他の方法で作り返されてもよい。新たな学習モデルMが作成されるまでは、現状の学習モデルMで不正検知が実行される。新たな学習モデルMが作成されると、新たな学習モデルMで不正検知が実行される。

【0106】

第2実施形態によれば、認証済み情報に基づいて、学習モデルMからの出力を取得し、認証済み情報に対応する出力に基づいて、学習モデルMの精度を評価する。認証済みユーザが正当である確率が非常に高いことに着目することによって、学習モデルMの精度を正確に評価できる。例えば、あるユーザの行動が正当であるか不正であるかを人手で判定しにくいことがある。更に、人手で判定できたとしても時間を要することがある。この点、認証済みユーザが正当であるとみなすことによって、学習モデルMの精度を迅速に評価できる。学習モデルMの精度が低下したことを迅速に検知し、最近の傾向に迅速に対応できるので、サービスにおける不正利用を防止し、セキュリティが高まる。本来は正当なはずの対象ユーザの行動が不正と推定されてしまいサービスを利用できなくなる、といった利便性の低下を防止できる。

20

30

【0107】

また、不正検知システムSは、複数の認証済み情報の各々に対応する出力を取得し、複数の認証済み情報の各々に対応する出力に基づいて、学習モデルMの精度を評価することによって、学習モデルMの精度をより正確に評価できる。学習モデルMの精度が低下したことを、より迅速に検知できる。学習モデルMの精度が低下したことを迅速に検知し、最近の傾向に迅速に対応できるので、サービスにおける不正利用をより確実に防止し、セキュリティが効果的に高まる。本来は正当なはずの対象ユーザの行動が不正と推定されてしまいサービスを利用できなくなる、といったこともより確実に防止できる。

【0108】

また、不正検知システムSは、所持認証を実行した認証済みユーザの認証済み情報を利用して学習モデルMを評価することによって、正当である確率が非常に高い認証済みユーザの認証済み情報を利用して、学習モデルMの精度をより正確に評価できる。学習モデルMの精度が低下したことを迅速に検知し、最近の傾向に迅速に対応できるので、サービスにおける不正利用をより確実に防止し、セキュリティが効果的に高まる。本来は正当なはずの対象ユーザの行動が不正と推定されてしまいサービスを利用できなくなる、といった利便性の低下もより確実に防止できる。

40

【0109】

【3. 変形例】

なお、本開示は、以上に説明した実施の形態に限定されるものではない。本開示の趣旨を逸脱しない範囲で、適宜変更可能である。

50

【0110】

[3-1. 第1実施形態に係る変形例]

まず、第1実施形態に係る変形例を説明する。

【0111】

[変形例1-1]

例えば、不正検知システムSは、任意のサービスに適用可能である。変形例1-1では、ユーザ端末20から利用可能な電子決済サービスに不正検知システムSを適用した場合を例に挙げる。変形例1-1以外の第1実施形態に係る変形例（変形例1-2～変形例1-10）と、第2実施形態に係る変形例（変形例2-1～2-9）と、も同様に、電子決済サービスを例に挙げる。

10

【0112】

電子決済サービスは、所定の決済手段を利用して電子決済を実行するサービスである。ユーザは、種々の決済手段を利用可能である。例えば、決済手段は、クレジットカード、デビットカード、電子マネー、電子キャッシュ、ポイント、銀行口座、ウォレット、又は仮想通貨であってもよい。バーコード又は二次元コード等のコードを利用した電子決済も、コード決済と呼ばれることがあるので、コードが決済手段に相当してもよい。

【0113】

変形例1-1における認証は、ユーザ端末20から実行される、電子決済サービスの認証である。認証済み情報は、電子決済サービスにおける認証済みユーザの行動に関する情報である。学習モデルMは、電子決済サービスにおける不正を検知するためのモデルである。以降、電子決済サービスを、単にサービスと記載する。

20

【0114】

変形例1-1の不正検知システムSは、ユーザのカードを利用したサービスを提供する。カードの一例として、クレジットカードを説明する。カードは、電子決済で利用可能なカードであればよく、クレジットカードに限られない。例えば、カードは、デビットカード、ポイントカード、電子マネーカード、キャッシュカード、交通系カード、又はその他の任意のカードであってもよい。カードは、ICカードに限られず、ICチップを含まないカードであってもよい。例えば、カードは、磁気カードであってもよい。

【0115】

図13は、変形例1-1の不正検知システムSの全体構成の一例を示す図である。不正検知システムSは、図1と同様の全体構成であってもよいが、変形例1-1では、他の全体構成の一例を説明する。図13に示すように、変形例の不正検知システムSは、ユーザ端末20、事業者サーバ30、及び発行者サーバ40を含む。不正検知システムSは、少なくとも1つのコンピュータを含めばよく、図13の例に限られない。ユーザ端末20、事業者サーバ30、及び発行者サーバ40の各々は、ネットワークNに接続される。ユーザ端末20は、第1実施形態及び第2実施形態と同様である。

30

【0116】

事業者サーバ30は、サービスを提供する事業者のサーバコンピュータである。事業者サーバ30は、制御部31、記憶部32、及び通信部33を含む。制御部31、記憶部32、及び通信部33の物理的構成は、それぞれ制御部11、記憶部12、及び通信部13と同様である。

40

【0117】

発行者サーバ40は、クレジットカードを発行した発行者のサーバコンピュータである。発行者は、事業者と同じであってもよいが、変形例1-1では、発行者が事業者とは異なる場合を説明する。発行者及び事業者は、互いに連携可能なグループ会社であってもよい。発行者サーバ40は、制御部41、記憶部42、及び通信部43を含む。制御部41、記憶部42、及び通信部43の物理的構成は、それぞれ制御部11、記憶部12、及び通信部13と同様である。

【0118】

なお、記憶部32、42に記憶されるプログラム及びデータの少なくとも一方は、ネッ

50

トワークNを介して供給されてもよい。また、事業者サーバ30及び発行者サーバ40の少なくとも一方に、コンピュータ読み取り可能な情報記憶媒体を読み取る読取部（例えば、光ディスクドライブやメモ리카ードスロット）と、外部機器とデータの入出力をするための入出力部（例えば、USBポート）と、の少なくとも一方が含まれてもよい。例えば、情報記憶媒体に記憶されたプログラム及びデータの少なくとも一方が、読取部及び入出力部の少なくとも一方を介して供給されてもよい。

【0119】

変形例1-1では、ユーザ端末20に、電子決済用のアプリケーション（以降、単にアプリ）がインストールされている。ユーザは、予め利用登録を済ませており、ユーザID及びパスワードでサービスにログインできるものとする。ユーザは、アプリから任意の決済手段を利用できる。変形例1-1では、ユーザがアプリからクレジットカード及び電子キャッシュを利用する場合を例に挙げる。以降、クレジットカードを、単にカードと記載する。

【0120】

図14は、変形例1-1のユーザ端末20に表示される画面の一例を示す図である。図14に示すように、ユーザがユーザ端末20を操作してアプリを起動させると、アプリのトップ画面G9が表示部25に表示される。トップ画面G9には、電子決済用のコードC90が表示される。例えば、店舗のPOS端末やコードリーダーでコードC90が読み取られると、予め設定された支払元の決済手段に基づいて、決済処理が実行される。コードC90を利用した決済処理自体は、公知の方法を利用可能である。

【0121】

図14の例では、「カード1」の名前で登録されたカードが支払元として設定されている。この状態でコードC90が読み取られると、このカードを利用した決済処理が実行される。ユーザは、支払元として設定したカードを利用して、アプリの電子キャッシュにチャージすることもできる。電子キャッシュは、オンライン上の電子マネーである。ユーザが支払元を電子キャッシュに変更してコードC90が読み取られると、電子キャッシュを利用した決済処理が実行される。

【0122】

変形例1-1では、トップ画面G9から、新たなカードを登録できるようになっている。例えば、ユーザがボタンB91を選択すると、新たなカードを登録するための登録画面G10が表示部25に表示される。ユーザは、入力フォームF100からカード番号、有効期限、及び名義人といったカード情報を入力する。変形例1-1では、カードの登録時の認証として、NFC認証、画像認証、及びセキュリティコード認証といった複数の認証方法が用意されている。ユーザは、ボタンB101～B103を選択し、任意の認証方法を選択できる。なお、クレジットカードの登録時の認証は、他の認証方法であってもよく、例えば、3Dセキュアと呼ばれる認証方法が利用されてもよい。

【0123】

NFC認証は、第1実施形態及び第2実施形態と同様であり、カードをNFC部23Aで読み取ることによって実行される。画像認証も、第1実施形態及び第2実施形態と同様であり、カードを撮影部26で撮影することによって実行される。セキュリティコード認証は、カードの裏面に形成されたセキュリティコードを操作部24から入力することによって実行される。セキュリティコードは、原則としてカードを所持していないと分からない情報なので、変形例1-1では、NFC認証及び画像認証だけではなく、セキュリティコード認証も所持認証の一例として説明する。

【0124】

図14では、セキュリティコード認証の流れが示されている。例えば、ユーザがボタンB103を選択すると、セキュリティコード認証を実行するための認証画面G11が表示部25に表示される。ユーザが、入力フォームF110にセキュリティコードを入力してボタンB111を選択すると、ユーザ端末20は、事業者サーバ30に、入力フォームF100に入力されたカード情報と、入力フォームF110に入力されたセキュリティコー

ドと、を送信する。以降、これらカード情報及びセキュリティコードを、それぞれ入力カード情報及び入力セキュリティコードと記載する。

【0125】

事業者サーバ30は、ユーザ端末20から入力カード情報及び入力セキュリティコードを受信すると発行者サーバ40に転送し、発行者サーバ40によりセキュリティコード認証が実行される。以降、発行者サーバ40に予め登録されているカード情報及びセキュリティコードを、それぞれ登録カード情報及び登録セキュリティコードと記載する。入力カード情報及び入力セキュリティコードの組み合わせと同じ登録カード情報及び登録セキュリティコードの組み合わせが発行者サーバ40に存在する場合に、セキュリティコード認証が成功する。

10

【0126】

セキュリティコード認証が実行されると、入力フォームF100から入力カード情報を入力したカードの登録が完了する。ユーザ端末20には、カードの登録が完了したことを示す完了画面G12が表示部25に表示される。以降、ユーザは、登録が完了したカードを支払元として設定できるようになる。

【0127】

変形例1-1では、個々のカードに、アプリから利用可能な上限額が設定される。この上限額は、カード自体の上限額（いわゆる利用枠又は限度額）を意味してもよいが、変形例1-1では、カード自体の上限額ではなく、アプリにおける上限額であるものとする。例えば、上限額は、所定期間（例えば、1週間又は1月）にアプリから利用可能な合計額である。上限額は、決済処理の1回あたりの上限額であってもよい。

20

【0128】

カードの上限額は、カードの登録時に実行された所持認証の認証方法によって異なる。カードの登録時に実行された所持認証のセキュリティが高いほど、このカードの上限額が高くなる。例えば、セキュリティコードは、フィッシング等によって流出することがあるので、セキュリティコード認証は、セキュリティが最も低いものとする。一方、NFC認証又は画像認証は、原則として、物理的なカードを所持していなければ成功させることができないので、セキュリティコード認証よりもセキュリティが高いものとする。

【0129】

図14の例では、セキュリティが最も低いセキュリティコード認証が実行されたので、上限額は、最も低い3万円となっている。例えば、ユーザがカードの登録時にボタンB101又はボタンB102を選択してNFC認証又は画像認証を実行すれば、上限額は、3万円よりも高い10万円になる。ユーザは、カードの登録後に、セキュリティが高い認証方法の所持認証を実行し、上限額を増やすこともできる。

30

【0130】

図15は、カードの登録後に上限額を増やす流れの一例を示す図である。図14のトップ画面G9のボタンB92が選択されると、図15に示すように、所持認証を実行するカードを選択するための選択画面G13が表示部25に表示される。選択画面G13には、登録済みのカードのリストL130が表示される。ユーザは、リストL130の中から所持認証を実行するカードを選択する。

40

【0131】

ユーザは、任意の認証方法を選択できる。例えば、ユーザがセキュリティコード認証を実行したカードを選択すると、ユーザは、セキュリティコード認証よりもセキュリティの高いNFC認証又は画像認証を選択できる。ユーザがボタンB131を選択すると、読取画面G6と同様の読取画面G14が表示部25に表示される。読取画面G14が表示されると、ユーザは、自身が所持するカードにユーザ端末20を近づける。

【0132】

図16は、カードのICチップをNFC部23Aで読み取る様子の一例を示す図である。図16では、電子マネー機能付きのカードC2を例に挙げる。カードC2の電子マネーは、アプリから利用可能であってもよいが、変形例1-1では、カードC2の電子マネー

50

は、アプリから利用できないものとする。即ち、カードC2の電子マネーは、アプリから利用可能な電子キャッシュとは異なる。カードC2の電子マネーは、所持認証で利用される。即ち、変形例1-1では、アプリで提供されるサービスとは直接的に関係のない他のサービスにおける電子マネーを利用して、所持認証が実行される。

【0133】

ICチップcpには、電子マネーを識別可能な電子マネーIDが記録されている。図16に示すように、ユーザがカードC2のICチップcpにユーザ端末20を近づけると、NFC部23Aは、ICチップcpに記録された情報を読み取る。NFC部23Aは、ICチップcp内の任意の情報を読み取り可能である。変形例1-1では、NFC部23Aが、ICチップcpに記録された電子マネーIDを読み取る場合を説明する。

10

【0134】

ユーザ端末20は、事業者サーバ30に、ICチップcpから読み取った電子マネーIDを送信する。この電子マネーIDは、ユーザ端末20から事業者サーバ30に入力されるので、以降では、この電子マネーIDを入力電子マネーIDと記載する。発行者サーバ40には、正解となる電子マネーIDが登録されている。以降、この電子マネーIDを登録電子マネーIDと記載する。なお、入力電子マネーIDと登録電子マネーIDを特に区別しないときは、単に電子マネーIDと記載することがある。

【0135】

事業者サーバ30は、発行者サーバ40に、ユーザ端末20から受信した入力電子マネーIDを転送する。その際に、ユーザがリストL130から選択したカードC2の入力カード情報も送信されるものとする。ユーザがカードC2の正当な持ち主であれば、入力カード情報及び入力電子マネーIDの組み合わせと同じ登録カード情報及び登録電子マネーIDの組み合わせが発行者サーバ40に登録されている。

20

【0136】

入力カード情報及び入力電子マネーIDの組み合わせと同じ登録カード情報及び登録電子マネーIDの組み合わせが発行者サーバ40に登録されている場合、所持認証が成功する。この場合、所持認証が成功したことを示す成功画面G15が表示部25に表示される。成功画面G15のように、NFC認証が実行されると、カードC2（図15の例では「カード2」）の上限額が3万円から10万円に増える。

【0137】

変形例1-1では、NFC認証が実行されたカードC2とは異なる他のカード（図15の例では「カード1」）の上限額も3万円から10万円に増えるものとするが、他のカードの上限額は増えなくてもよい。なお、NFC認証が実行されたカードC2と同じユーザIDに関連付けられていたとしても、名義人が異なれば、第三者が勝手に登録した可能性があるので、上限額は増えないものとする。入力カード情報及び入力電子マネーIDの組み合わせと同じ登録カード情報及び登録電子マネーIDの組み合わせが発行者サーバ40に登録されていない場合、所持認証が失敗する。この場合、図3の失敗画面G8と同様の失敗画面G16が表示部25に表示される。

30

【0138】

画像認証も同様の流れで実行される。NFC認証ではNFC部23Aを利用して入力電子マネーIDが取得されるのに対し、画像認証では、カードC2が撮影された撮影画像を利用して入力電子マネーIDが取得される。例えば、ユーザが選択画面G13のボタンB132を選択すると、撮影部26が起動する。撮影部26は、カードC2を撮影する。図16のカードC2の例では、裏面に入力電子マネーIDが形成されているものとするが、表面に入力電子マネーIDが形成されていてもよい。

40

【0139】

ユーザがカードC2の裏面を撮影すると、ユーザ端末20は、事業者サーバ30に、撮影画像を送信する。事業者サーバ30は、撮影画像を受信すると、撮影画像に光学文字認識を実行して入力カード情報を取得する。入力カード情報が取得された後の流れは、NFC認証と同様である。光学文字認識は、ユーザ端末20で実行されてもよい。第1実施形

50

態の入力個人番号と同様、入力電子マネーIDは、バーコード又は二次元コード等のコードに含まれていてもよい。

【0140】

なお、所持認証で利用される情報は、入力電子マネーIDに限られない。例えば、カードC2がポイントカードの機能も有している場合には、ポイントを識別可能なポイントIDが所持認証で利用されてもよい。ポイントIDは、カードC2に含まれているものとする。他にも例えば、カードC2のカード番号や有効期限が所持認証で利用されてもよい。変形例1-1では、カードC2に含まれる何らかの情報又はこの情報に関連付けられた情報が所持認証で利用されるようにすればよく、カードC2のデザインや発行日等が所持認証で利用されてもよい。

10

【0141】

図17は、第1実施形態に係る変形例における機能ブロック図である。図17では、変形例1-1以降の変形例1-2～1-10における機能についても示している。図17に示すように、ここでは、事業者サーバ30により主な機能が実現される場合を説明する。事業者サーバ30では、データ記憶部300、認証済み情報取得部301、作成部302、不正検知部303、比較部304、未認証情報取得部305、及び確定情報取得部306が実現される。データ記憶部300は、記憶部32を主として実現される。他の各機能は、制御部31を主として実現される。

【0142】

データ記憶部300は、データ記憶部300は、ユーザデータベースDB1、訓練データベースDB2、及び学習モデルMを記憶する。これらのデータは、概ね第1実施形態と同様であるが、ユーザデータベースDB1の具体的な内容が第1実施形態とは異なる。

20

【0143】

図18は、ユーザデータベースDB1のデータ格納例を示す図である。図18に示すように、ユーザデータベースDB1は、利用登録が完了したユーザに関する情報が格納されたデータベースである。例えば、ユーザデータベースDB1には、ユーザID、パスワード、氏名、支払元の決済手段、登録カード情報、電子キャッシュ情報、場所情報、日時情報、及び利用情報が格納される。例えば、ユーザが利用登録をすると、ユーザIDが発行され、ユーザデータベースDB1に新たなレコードが作成される。このレコードには、利用登録時に指定されたパスワード及び氏名とともに、登録カード情報及び電子キャッシュ情報が格納される。

30

【0144】

登録カード情報は、ユーザが登録したカードC2に関する情報である。例えば、登録カード情報は、個々のユーザの中でカードを識別するための連番の数値、カード番号、有効期限、名義人、所持認証フラグ、及び利用設定を含む。先述したように、変形例1-1の利用設定は、アプリから利用可能なカードC2の上限額の設定である。ユーザが新たなカードC2を登録すると、このカードC2に対応する登録カード情報が追加される。

【0145】

電子キャッシュ情報は、アプリから利用可能な電子キャッシュに関する情報である。例えば、電子キャッシュ情報は、電子キャッシュを識別可能な電子キャッシュIDと、電子キャッシュの残高と、を含む。電子キャッシュは、ユーザが登録したカードC2でチャージ可能であってもよい。この場合のチャージの上限額の設定が利用設定に相当してもよい。なお、ユーザデータベースDB1に格納される情報は、図18の例に限られない。

40

【0146】

場所情報、日時情報、及び利用情報の組み合わせが認証済み情報に相当する点は、第1実施形態と同様である。変形例では、場所情報は、決済処理が実行された場所を示す。この場所は、店舗や自動販売機等が配置された場所である。日時情報は、決済処理が実行された日時である。利用情報は、利用額、購入された商品、利用した決済手段（決済処理の実行時に設定されていた支払元の決済手段）等の情報である。図18のデータ格納例では、ユーザID及び端末IDの組み合わせごとに、場所情報、日時情報、及び利用情報が格

50

納されるが、場所情報、日時情報、及び利用情報は、ユーザIDごとやカードC2ごとに格納されてもよい。

【0147】

認証済み情報取得部301、作成部302、及び不正検知部303は、それぞれ認証済み情報取得部101、作成部102、及び不正検知部103と同様である。変形例1-1における学習モデルMは、不正な決済処理を検知するためのモデルである。作成部302は、認証済みユーザが決済処理を実行した店舗等の場所情報、決済処理を実行した日時情報、及び決済額等の利用情報を入力した場合に、正当であることを示す情報が出力されるように、学習モデルMを作成する。

【0148】

不正検知部103は、対象ユーザが決済処理を実行した店舗等の場所情報、決済処理を実行した日時情報、及び決済額等の利用情報に基づいて、学習モデルMからの出力を取得し、当該出力が不正を示すか否かを判定することによって、不正を検知する。例えば、変形例1-1における不正は、第三者による不正なログインによる決済手段の利用行為、第三者が不正に入手したカード番号を自身のユーザIDに登録して店舗における決済処理を実行する行為、又は第三者が不正に入手したカード番号を利用して自身の電子マネー若しくは電子キャッシュにチャージする行為等である。第三者が不正にログインして支払元を変更する行為、登録カード情報を勝手に登録する行為、又はその他の設定や登録情報を変更する行為は、不正に相当する。

【0149】

変形例1-1によれば、決済における不正を検知するための学習モデルMの作成を簡易化できる。

【0150】

[変形例1-2]

例えば、変形例1-1のようなサービスでは、認証済みユーザは、所定のカードC2である第1カードC2と、第2カードC3と、の各々を利用可能であってもよい。変形例1-2では、第1カードC2は、所持認証が実行されるカードである場合を説明するが、第1カードC2の認証方法は、所持認証に限られない。第1カードC2の認証方法は、任意の認証方法であってもよく、例えば、知識認証又は生体認証であってもよい。3Dセキュアは、知識認証の一例である。他の認証方法の例は、第1実施形態で説明した通りである。第1カードC2は、先述の所定の認証が実行されるカードであればよい。

【0151】

変形例1-2では、第1カードC2と区別するために、第2カードにC3の符号を付すが、第2カードC3は、図面には示していない。第1カードC2に関連付けられた第2カードC3とは、第1カードC2と同じユーザIDに関連付けられた第2カードC3である。ユーザIDを介するのではなく、第1カードC2と第2カードC3が直接的に関連付けられてもよい。

【0152】

第2カードC3は、所持認証が実行されていないカードである。第2カードC3は、所持認証を実行可能ではあるが、単に所持認証が実行されていないカードであってもよい。第2カードC3が所持認証を実行可能なカードである場合には、第2カードC3が第1カードC2に相当することもある。変形例1-2では、第2カードC3は、NFC認証又は画像認証に対応していないカードである。例えば、第2カードC3は、NFC認証又は画像認証で利用される入力電子マネーIDを含まない。

【0153】

例えば、第2カードC3がICチップを含んでいたとしても、このICチップには、入力電子マネーIDは含まない。このICチップに、何らかの電子マネーIDが含まれていたとしても、NFC認証又は画像認証では利用されない他の電子マネーの電子マネーIDである。同様に、第2カードC3に何らかの電子マネーIDが形成されていたとしても、NFC認証又は画像認証では利用されない他の電子マネーの電子マネーIDである。

【0154】

認証済み情報取得部101は、第1カードC2に対応する認証済み情報を取得する。この認証済み情報は、所持認証フラグが「1」又は「2」である第1カードC2の認証済み情報である。認証済み情報取得部101は、ユーザデータベースDB1を参照し、利用情報が示す決済手段が第1カードC2であると共に、所持認証フラグが「1」又は「2」であるレコードを特定し、このレコードに格納された場所情報、日時情報、及び利用情報を、認証済み情報として取得する。

【0155】

作成部302は、第1カードC2に対応する認証済み情報に基づいて、学習モデルMを作成する。作成部302は、第2カードC3に対応する場所情報、日時情報、及び利用情報は、学習モデルMの作成で利用しなくてもよい。認証済み情報に基づいて学習モデルMを作成する方法自体は、第1実施形態で説明した通りである。

10

【0156】

変形例1-2によれば、第1カードC2に対応する認証済み情報に基づいて、学習モデルMを作成する。正当である確率が非常に高い第1カードC2に対応する認証済み情報に着目することによって、第1実施形態で説明した学習モデルMの作成を簡易化、学習モデルMの迅速な作成、サービスにおける不正利用の防止、セキュリティの向上、及び利便性の低下の防止を効果的に実現できる。

【0157】

[変形例1-3]

例えば、第2カードC3の所持認証が実行されていないとしても、所持認証が実行された第1カードC2と同じ名義人であれば、第2カードC3を利用した行動も正当である確率が非常に高い。このため、名義人が同じことを条件として、第2カードC3の場所情報、日時情報、及び利用情報を、認証済み情報として利用してもよい。

20

【0158】

不正検知システムSは、第1カードC2の名義に関する第1名義情報と、第2カードC3の名義に関する第2名義情報と、を比較する比較部304を更に含む。第1名義情報は、第1カードC2の名義に関する情報である。第2名義情報は、第2カードC3の名義に関する情報である。変形例1-3では、第1名義情報が第1カードC2の名義人である第1名義人を示し、第2名義情報が第2カードC3の名義人である第2名義人を示す場合を説明する。

30

【0159】

第1名義人は、第1カードC2の名義人の名前を示す文字列である。第2名義人は、第2カードC3の名義人の名前を示す文字列である。名義人は、任意の言語の文字列で表現可能である。なお、第1名義情報及び第2名義情報の各々は、名義人以外の情報であってもよい。例えば、第1名義情報及び第2名義情報の各々は、名義人の住所、電話番号、生年月日、性別、メールアドレス、又はこれらの組み合わせであってもよいし、他の個人情報であってもよい。

【0160】

変形例1-3では、比較部304が事業者サーバ30により実現される場合を説明するが、比較部304は、発行者サーバ40により実現されてもよい。例えば、ユーザデータベースDB1に格納されていない情報を第1名義情報及び第2名義情報として利用する場合には、第1名義情報及び第2名義情報の比較は、発行者サーバ40により実行されてもよい。ここでの比較とは、一致しているか否かを判定することである。

40

【0161】

変形例1-3では、データ記憶部300は、種々のカードに関する情報が格納されたデータベースを記憶するものとする。このデータベースには、種々のカードの名義情報が格納されているものとする。第1名義情報及び第2名義情報は、このデータベースから取得される。事業者サーバ30でこのデータベースを管理しない場合には、事業者サーバ30は、発行者サーバ40に第1名義情報及び第2名義情報の比較を依頼し、発行者サーバ4

50

0 から比較結果のみを取得すればよい。例えば、比較部 304 は、第 1 名義人と、第 2 名義人と、を比較する。比較部 304 は、ユーザデータベース DB1 を参照し、第 1 名義人及び第 2 名義人を取得し、これらの比較結果を認証済み情報取得部 101 に送る。第 1 名義情報及び第 2 名義情報が他の情報であってもよい点は、先述した通りである。

【0162】

認証済み情報取得部 101 は、比較部 304 の比較結果が所定の結果である場合に、第 2 カード C3 に対応する認証済み情報を取得する。変形例 1-3 では、第 1 名義人及び第 2 名義人が一致することが所定の結果に相当する場合を説明するが、先述した他の情報が一致することが所定の結果に相当してもよい。第 1 名義情報及び第 2 名義情報の各々に複数の情報が含まれる場合には、所定数以上の情報が一致することが所定の結果に相当してもよい。例えば、第 1 名義情報及び第 2 名義情報の各々に、名義人、住所、電話番号、及び生年月日といった 4 つの情報が含まれる場合に、2 つ以上の情報が一致することが所定の結果に相当してもよい。なお、ここでの一致とは、完全一致ではなく、部分一致であってもよい。

10

【0163】

図 18 の例であれば、ユーザ ID 「taro.yamada123」の第 1 カード C2 (No. 2 のカード) の第 1 名義人と、第 2 カード C3 (No. 1 のカード) の第 2 名義人と、は両方とも「TARO YAMADA」で同じである。このため、第 1 カード C2 の所持認証が実行されると、第 2 カード C3 も学習モデル M の学習で利用される。

【0164】

20

一方、ユーザ ID 「hanako.suzuki999」の第 1 カード C2 (No. 1 のカード) の第 1 名義人と、ある第 2 カード C3 (No. 2 のカード) の第 2 名義人と、は両方とも「HANAKO SUZUKI」で同じである。このため、第 1 カード C2 の所持認証が実行されると、この第 2 カード C3 も学習モデル M の学習で利用される。ただし、他の第 2 カード C3 (No. 3 のカード) の第 2 名義人は、「MIKI OKAMOTO」であり、第 1 名義人とは異なる。このため、当該他の第 2 カード C3 は、第三者が勝手に登録した可能性があり、当該他の第 2 カード C3 を利用した行動は、正当ではないかもしれないので、学習モデル M の学習で利用されない。

【0165】

作成部 302 は、比較部 304 の比較結果が所定の結果である場合に、第 1 カード C2 に対応する認証済み情報と、第 2 カード C3 に対応する認証済み情報と、に基づいて、学習モデル M を作成する。第 2 カード C3 は、所持認証が実行されていないため、第 2 カード C3 の場所情報、日時情報、及び利用情報は、厳密には認証済み情報には該当しないが、第 1 カード C2 に対応する認証済み情報と同等に扱うので、ここでは、第 2 カード C3 に対応する認証済み情報と記載する。第 2 カード C3 に対応する認証済み情報が学習で用いられる点が第 1 実施形態及び変形例 1-1 と異なるだけであり、学習モデル M の学習方法自体は、第 1 実施形態及び変形例 1-1 と同様である。作成部 302 は、第 1 カード C2 に対応する認証済み情報と、第 2 カード C3 に対応する認証済み情報と、の各々を学習モデル M に入力した場合に、正当と推定されるように、学習モデル M を作成する。

30

【0166】

40

変形例 1-3 によれば、第 1 カード C2 の名義に関する第 1 名義情報と、第 2 カード C3 の名義に関する第 2 名義情報と、の比較結果が所定の結果である場合に、第 1 カード C2 に対応する認証済み情報と、第 2 カード C3 に対応する認証済み情報と、に基づいて、学習モデル M を作成することによって、より多くの認証済み情報を学習させて学習モデル M の精度がより高まる。その結果、サービスにおける不正利用の防止、セキュリティの向上、及び利便性の低下の防止を効果的に実現できる。

【0167】

[変形例 1-4]

例えば、変形例 1-3 で説明した第 2 カード C3 は、所持認証に対応していないカードであってもよい。第 2 カード C3 に対応する認証済み情報は、所持認証が実行されてい

50

い第2カードC3を利用した認証済みユーザの行動に関する情報であってもよい。所持認証に対応していないカードは、所持認証を実行できないカードである。例えば、ICチップを含んでいないカードは、NFC認証に対応しない。例えば、券面に入力電子マネーIDが形成されていないカードは、画像認証に対応しない。例えば、所持認証で利用される入力電子マネーIDを含まないカードは、所持認証に対応していないカードである。

【0168】

変形例1-4によれば、第2カードC3が所持認証に対応していないカードであったとしても、第2カードC3に対応する認証済み情報に基づいて、学習モデルMを作成することによって、学習モデルMの精度がより高まる。

【0169】

[変形例1-5]

例えば、所持認証を実行していない未認証ユーザの行動を利用して学習モデルMの学習が行われてもよい。不正検知システムSは、認証を実行していない未認証ユーザの行動に関する未認証情報を取得する未認証情報取得部305を更に含む。未認証ユーザは、所持認証フラグが「1」又は「2」ではないユーザである。即ち、未認証ユーザは、所持認証フラグの少なくとも一部が「0」のユーザである。未認証情報取得部305は、ユーザデータベースDB1を参照し、未認証ユーザの未認証情報を取得する。未認証情報は、未認証ユーザの場所情報、日時情報、及び利用情報の組み合わせである。未認証情報が任意の情報であってもよく、これらの組み合わせに限られない点は、認証済み情報と同様である。

【0170】

作成部302は、未認証情報に基づいて、未認証ユーザの行動が正当又は不正であることを示す訓練データを作成し、この訓練データに基づいて、学習モデルMを学習させる。以降、認証済みユーザを利用して作成された訓練データを第1訓練データと記載し、未認証ユーザを利用して作成された訓練データを第2訓練データと記載する。第1訓練データ及び第2訓練データのデータ構造自体は、共に同じであり、第1実施形態で説明した通りである。

【0171】

なお、第1訓練データの出力部分は、原則として必ず正当を示すのに対し、第2訓練データの出力部分は、正当を示すとは限らない。例えば、第2訓練データの出力部分は、学習モデルMの作成者により指定される。学習モデルMの作成者により不正と判定された未認証ユーザについては、第2訓練データの出力部分は、不正を示す。第1訓練データ及び第2訓練データのデータ構造自体は、共に同じなので、第1訓練データ及び第2訓練データの各々に基づいて学習モデルMを作成する方法自体は、第1実施形態で説明した通りである。

【0172】

変形例1-5によれば、未認証情報に基づいて、未認証ユーザの行動が正当又は不正であることを示す第2訓練データを作成し、第2訓練データに基づいて、学習モデルMを学習させることによって、より多くの情報を利用して学習モデルMの精度がより高まる。

【0173】

[変形例1-6]

例えば、変形例1-5において、作成部302は、未認証情報に基づいて、学習済みの学習モデルMからの出力を取得し、当該出力に基づいて、第2訓練データを作成してもよい。例えば、作成部302は、未認証情報に対応する学習モデルMの出力を、学習モデルMの作成者に提示する。学習モデルMの作成者は、この出力が正しいか否かをチェックする。作成者は、必要に応じてこの出力を修正する。

【0174】

例えば、未認証ユーザが本来は正当と思われるが、学習モデルMからの出力が不正を示す場合に、作成者は、この出力を正に修正する。逆に、未認証ユーザが本来は不正と思われるが、学習モデルMからの出力が正当を示す場合に、作成者は、この出力を不正に修正する。作成部302は、未認証ユーザの修正結果に基づいて、第2訓練データを作成す

10

20

30

40

50

る。作成部 302 は、未認証ユーザが出力を修正しなかった場合には、学習モデル M からの出力に基づいて、第 2 訓練データを作成する。第 2 訓練データを利用して学習モデル M を作成する方法自体は、変形例 1-5 で説明した通りである。

【0175】

変形例 1-6 によれば、未認証情報に基づいて、学習済みの学習モデル M からの出力を取得し、当該出力に基づいて、第 2 訓練データを作成することによって、より多くの情報を利用して学習モデル M の精度がより高まる。

【0176】

[変形例 1-7]

例えば、変形例 1-5 において、ある未認証ユーザがサービスを継続して利用している間に、この未認証ユーザが不正であるか正当であるかが徐々に分かっていくことがある。このため、作成部 302 は、未認証情報に対応する出力が取得された後の未認証情報に基づいて、当該出力の内容を変更し、当該変更された出力の内容に基づいて、第 2 訓練データを作成してもよい。

【0177】

変形例 1-7 の学習モデル M は、サービスにおける不正に関するスコアを出力するものとする。変形例 1-7 では、スコアが正当の度合いを示す場合を説明するが、スコアは不正の度合いを示してもよい。スコアが正当の度合いを示す場合、スコアは、正当に分類される蓋然性を示す。スコアが不正の度合いを示す場合、スコアは、不正に分類される蓋然性を示す。学習モデル M がスコアを計算する方法自体は、公知の種々の方法を利用可能である。作成部 302 は、未認証ユーザの未認証行動に基づいて、学習モデル M からのスコアを取得する。作成部 302 は、その後の未認証ユーザの行動に基づいて、このスコアを変更する。スコアの変更方法は、予めデータ記憶部 100 に定められているものとする。

【0178】

例えば、不正に分類される行動と、この行動が行われた場合のスコアの変更量（本変形例では、スコアが正当の度合いを示すので減少量）と、の関係が定められている。同様に、正当に分類される行動と、この行動が行われた場合のスコアの変更量（本変形例では、スコアが正当の度合いを示すので増加量）と、の関係が定められている。作成部 302 は、未認証ユーザが不正の疑いのある行動をした場合、不正の度合いが強くなるように、この行動に応じた変更量に基づいて、スコアを変更する。作成部 302 は、未認証ユーザが正当の疑いのある行動をした場合、不正の度合いが弱くなるように、この行動に応じた変更量に基づいて、スコアを変更する。

【0179】

なお、学習モデル M がスコアを出力するのではなく、不正であるか否かの分類結果を出力する場合には、作成部 302 は、この分類結果を変更してもよい。例えば、学習モデル M の出力が、不正であることを示す「1」、又は、正当であることを示す「0」だったとする。作成部 302 は、未認証情報に対応する出力が「1」であり、未認証ユーザが不正に分類された場合に、その後の未認証ユーザが正当の確率が高い行動を継続して行った場合には、この出力を「0」に変更したうえで、第 2 訓練データを作成してもよい。作成部 302 は、未認証情報に対応する出力が「0」であり、未認証ユーザが正当に分類された場合に、その後の未認証ユーザが不正の確率が高い行動を継続して行った場合には、この出力を「1」に変更したうえで、第 2 訓練データを作成してもよい。

【0180】

変形例 1-7 によれば、未認証情報に対応する出力が取得された後の未認証情報に基づいて、当該出力の内容を変更し、当該変更された出力の内容に基づいて、第 2 訓練データを作成することによって、学習モデル M の精度がより高まる。

【0181】

[変形例 1-8]

例えば、変形例 1-7 において、未認証情報に対応するスコアは、認証済み情報に対応するスコアよりも不正を示すように、上限値が設定されてもよい。作成部 302 は、学習

10

20

30

40

50

モデルMから出力された認証済み情報のスコアに基づいて、未認証情報に対応するスコアの上限值を決定する。例えば、作成部302は、認証済み情報のスコアの平均値を、未認証情報に対応するスコアの上限值として決定する。他にも例えば、作成部302は、認証済み情報のスコアのうちの最も低い値、又は、所定番目に低い値を、未認証情報に対応するスコアの上限值として決定する。

【0182】

学習モデルMは、上限値に基づいて、未認証情報に対応するスコアを出力する。学習モデルMは、上限値を超えないように、未認証情報に対応するスコアを出力する。例えば、学習モデルMの内部的に計算されたスコアが上限値を超えたとしても、学習モデルMは、出力されるスコアが上限値以下になるように、スコアを出力する。上限値は、未認証情報を学習モデルMに入力することによって得られたスコアの平均値等であってもよい。未認証情報に対応するスコアを利用して学習モデルMを作成する方法自体は、変形例1-7で説明した通りである。

【0183】

変形例1-8によれば、認証済み情報に対応するスコアよりも不正を示すように設定された上限値に基づいて、未認証情報に対応するスコアを出力することによって、学習モデルMの精度がより高まる。

【0184】

[変形例1-9]

例えば、所定の時間が経過して不正であるか否かが確定した確定ユーザの行動も利用して、学習モデルMを作成してもよい。不正検知システムSは、不正であるか否かが確定した確定ユーザの行動に関する確定情報を取得する確定情報取得部306を更に含む。確定情報は、確定ユーザの行動に関する情報であるという点で認証済み情報とは異なるが、データ構造自体は、認証済み情報と同様である。このため、確定情報は、ユーザデータベースDB1に格納された確定ユーザの場所情報、日時情報、及び利用情報を含む。確定情報に含まれる内容がこれらに限られない点も認証済み情報と同様である。不正であるか否かは、学習モデルMの作成者により指定されてもよいし、所定のルールに基づいて決定されてもよい。

【0185】

作成部302は、認証済み情報及び確定情報に基づいて、学習モデルMを作成する。確定情報が利用される点で第1実施形態及び他の変形例と異なるだけであり、学習モデルMの作成方法自体は、第1実施形態及び他の変形例と同様である。即ち、作成部302は、認証済み情報が入力された場合に、正当との結果が出力されるように、かつ、確定情報の各々が入力された場合に、確定情報に関連付けられた結果（不正であるか正当であるかの結果）が出力されるように、学習モデルMを作成する。

【0186】

変形例1-9によれば、認証済み情報と、確定ユーザの確定情報と、に基づいて、学習モデルMを作成することによって、より多くの情報を利用して学習させ、学習モデルMの精度がより高まる。

【0187】

[変形例1-10]

例えば、学習モデルMは、教師無し学習のモデルであってもよい。作成部302は、認証済み情報に基づいて、サービスにおける不正な行動が外れ値となるように、学習モデルMを作成する。例えば、作成部302は、複数の認証済み情報の各々が入力された場合に、これら認証済み情報が同じクラスタにクラスタリングされるように、教師無し学習の学習モデルMを作成する。この学習モデルMでは、認証済み情報が示す特徴とは異なる不正な行動に関する情報が入力されると、外れ値として出力される。即ち、不正な行動は、認証済み情報のクラスタに属さないものとして出力される。教師無し学習自体は、種々の方法を利用可能であり、例えば、上記のクラスタリングの方法以外にも、主成分分析、ベクトル量子化、非負値行列因子分解、k-means法、又は混合ガウスモデル等の方法を

10

20

30

40

50

利用可能である。不正検知部 303 は、対象ユーザの対象情報に基づいて、学習モデル M の出力を取得し、出力が外れ値であれば不正と判定する。不正検知部 303 は、出力が外れ値でなければ、正当と判定する。

【0188】

変形例 1-10 によれば、認証済み情報に基づいて、サービスにおける不正な行動が外れ値となるように、教師無し学習を利用した学習モデル M を作成することによって、教師無し学習を利用した学習モデル M の作成を簡易化できる。また、学習モデル M の作成の一連の処理を自動化し、学習モデル M を迅速に作成できる。不正検知システム S に、最新の傾向を学習させた学習モデル M を迅速に適用し、精度良く不正を検知できる。その結果、サービスにおける不正利用を防止し、セキュリティが高まる。本来は正当なはずの対象ユーザの行動が不正と推定されてしまいサービスを利用できなくなる、といった利便性の低下も防止できる。

【0189】

[3-2. 第 2 実施形態に係る変形例]

次に、第 2 実施形態に係る変形例を説明する。

【0190】

[変形例 2-1]

例えば、第 2 実施形態の不正検知システム S も、第 1 実施形態の変形例 1-1 ~ 変形例 1-10 で説明したような電子決済サービスに適用可能である。

【0191】

図 19 は、第 2 実施形態に係る変形例における機能ブロック図である。図 19 では、変形例 2-1 以降の変形例 2-2 ~ 2-9 における機能についても示している。図 19 に示すように、ここでは、事業者サーバ 30 により主な機能が実現される場合を説明する。事業者サーバ 30 では、データ記憶部 300、認証済み情報取得部 301、作成部 302、不正検知部 303、比較部 304、未認証情報取得部 305、確定情報取得部 306、出力取得部 307、評価部 308、及び処理実行部 309 を含む。出力取得部 307、評価部 308、及び処理実行部 309 の各々は、制御部 31 を主として実現される。

【0192】

データ記憶部 300、変形例 1-1 と同様である。認証済み情報取得部 301、不正検知部 303、及び評価部 308 は、第 2 実施形態で説明した認証済み情報取得部 301、不正検知部 303、及び評価部 308 と同様である。認証済み情報取得部 301 及び不正検知部 303 は、変形例 1-1 の認証済み情報取得部 301 及び不正検知部 303 とも共通した機能を有する。評価部 308 は、変形例 1-1 で説明したような第三者による不正なログインによる決済手段の利用等の不正を検知するための学習モデル M の正解率等を利用して、学習モデル M の精度を評価する。この評価の指標が正解率に限られないことは、第 2 実施形態で説明した通りである。

【0193】

変形例 2-1 によれば、電子決済サービスにおける不正を検知するための学習モデル M の不正検知の精度を正確に評価できる。

【0194】

[変形例 2-2]

例えば、不正検知システム S は、学習モデル M の精度が所定の精度未満になった場合に、サービスにおける最近の行動を利用して学習モデル M を作成するための処理を実行する処理実行部 309 を含んでもよい。この処理は、学習モデル M の作成者に学習モデル M を作成し直すように通知する処理であってもよいし、第 1 実施形態と同様の方法によって、学習モデル M を作成し直す処理であってもよい。第 2 実施形態で説明した通り、通知は、電子メール等の任意の手段を利用可能である。学習モデル M を作成し直す処理は、最近の認証済み情報を利用して、第 1 実施形態のように学習モデル M を作成する処理であってもよいし、特に第 1 実施形態のような学習モデル M の作成ではない方法が利用されてもよい。更に、学習モデル M は、不正検知システム S 以外のシステムで作成されてもよい。

【0195】

変形例2-2によれば、学習モデルMの精度が所定の精度未満になった場合に、サービスにおける最近の行動を利用して学習モデルMを作成するための処理を実行することによって、学習モデルMの不正検知の精度が低下した場合に対処できる。不正検知システムSに、最新の傾向を学習させた学習モデルMを迅速に適用し、精度良く不正を検知できる。その結果、サービスにおける不正利用を防止し、セキュリティが高まる。本来は正当なはずの対象ユーザの行動が不正と推定されてしまいサービスを利用できなくなる、といった利便性の低下も防止できる。

【0196】

[変形例2-3]

例えば、評価部308は、認証済み情報及び確定情報に基づいて、学習モデルMの精度を評価してもよい。変形例2-3の不正検知システムSは、変形例1-9と同様の確定情報取得部306を含むものとする。確定情報が学習モデルMの精度の評価で利用される点で第2実施形態とは異なるが、学習モデルMの評価方法自体は、第2実施形態で説明した通りである。例えば、評価部308は、認証済み情報だけでなく、確定情報も利用して、正解率を計算する。評価部308は、確定情報を学習モデルMに入力することによって得られた出力が、確定情報に対応する出力（例えば、学習モデルMの作成者が指定した不正であるか否かの結果）を示すか否かを判定し、正解率を計算する。正解率以外の任意の指標が利用可能である点は、第2実施形態で説明した通りである。

【0197】

変形例2-3によれば、認証済み情報及び確定情報に基づいて、学習モデルMの精度を評価することによって、より多くの情報を利用して学習モデルMの精度をより正確に評価できる。

【0198】

[変形例2-4]

例えば、変形例1-2と同様に、第1カードC2及び第2カードC3の各々を利用可能である場合に、出力取得部307は、第1カードC2に対応する認証済み情報に基づいて、第1カードC2に対応する出力を取得してもよい。評価部308は、第1カードC2に対応する出力に基づいて、学習モデルMの精度を評価する。学習モデルMの出力に基づいて学習モデルMの精度を評価する方法自体は、第2実施形態で説明した通りである。

【0199】

変形例2-4によれば、第1カードC2に対応する出力に基づいて、学習モデルMの精度を評価する。正当である確率が非常に高い第1カードC2に対応する認証済み情報に着目することによって、第2実施形態で説明した学習モデルMの正確な評価、最近の傾向への迅速な対応、サービスにおける不正利用の防止、セキュリティの向上、及び利便性の低下の防止を効果的に実現できる。

【0200】

[変形例2-5]

例えば、不正検知システムSが変形例1-3と同様の比較部304を含む場合に、出力取得部307は、第2カードC3に対応する認証済み情報に基づいて、第2カードC3に対応する出力を取得してもよい。評価部308は、第1カードC2に対応する出力と、第2カードC3に対応する出力と、に基づいて、学習モデルMの精度を評価する。学習モデルMの出力に基づいて学習モデルMの精度を評価する方法自体は、第2実施形態で説明した通りである。例えば、評価部308は、第1カードC2に対応する出力だけでなく、第2カードC3に対応する出力も利用して、正解率を計算する。評価部308は、第2カードC3に対応する認証済み情報を学習モデルMに入力することによって得られた出力が、正当を示すか否かを判定し、正解率を計算する。正解率以外の任意の指標が利用可能である点は、第2実施形態で説明した通りである。

【0201】

変形例2-5によれば、第1カードC2の名義に関する第1名義情報と、第2カードC

10

20

30

40

50

3の名義に関する第2名義情報と、の比較結果が所定の結果である場合に、第1カードC2に対応する出力と、第2カードC3に対応する出力と、に基づいて、学習モデルMの精度を評価することによって、より多くの情報を利用して学習モデルMをより正確に評価できる。その結果、サービスにおける不正利用の防止、セキュリティの向上、及び利便性の低下の防止を効果的に実現できる。

【0202】

[変形例2-6]

例えば、変形例1-4と同様に、変形例2-5の第2カードC3は、所持認証に対応していないカードであってもよい。変形例2-5で説明した第2カードC3が所持認証に対応していないだけであり、評価部308の評価方法自体は、変形例2-5で説明した通りである。

10

【0203】

変形例2-6によれば、第2カードC3が所持認証に対応していないカードであったとしても、第2カードC3に対応する認証済み情報に基づいて、学習モデルMの精度を評価することによって、より多くの情報を利用して学習モデルMをより正確に評価できる。

【0204】

[変形例2-7]

例えば、変形例1-1と同様に、不正検知システムSは、作成部302を含んでもよい。作成部302は、認証済み情報に基づいて、認証済みユーザの行動が正当と推定されるように、サービスにおける不正を検知するための学習モデルMを作成する。変形例2-7の不正検知システムSは、変形例1-1と同様の構成を有するようにすればよい。

20

【0205】

変形例2-7によれば、第1実施形態で説明した学習モデルMの作成を簡易化、学習モデルMの迅速な作成、サービスにおける不正利用の防止、セキュリティの向上、及び利便性の低下の防止を効果的に実現できる。

【0206】

[変形例2-8]

例えば、不正検知システムSは、変形例1-5と同様の未認証情報取得部305を含んでもよい。作成部302は、未認証情報に基づいて、未認証ユーザの行動が正当又は不正であることを示す第2訓練データを作成し、第2訓練データに基づいて、学習モデルMを学習させてもよい。変形例2-8の不正検知システムSは、変形例1-5と同様の構成を有するようにすればよい。更に、評価部308は、第2訓練データに基づいて作成した学習モデルMの精度を評価してもよい。この評価方法は、第2実施形態又は上記説明した変形例と同様の方法であればよい。

30

【0207】

変形例2-8によれば、未認証情報に基づいて、未認証ユーザの行動が正当又は不正であることを示す第2訓練データを作成し、第2訓練データに基づいて、学習モデルMを学習させることによって、より多くの情報を利用して学習モデルMの精度がより高まる。

【0208】

[変形例2-9]

例えば、変形例1-6と同様に、作成部302は、未認証情報に基づいて、学習済みの学習モデルMからの出力を取得し、当該出力に基づいて、第2訓練データを作成してもよい。変形例2-9の不正検知システムSは、変形例1-6と同様の構成を有するようにすればよい。

40

【0209】

変形例2-9によれば、未認証情報に基づいて、学習済みの学習モデルMからの出力を取得し、当該出力に基づいて、第2訓練データを作成することによって、より多くの情報を利用して学習モデルMの精度がより高まる。

【0210】

[3-3. その他の変形例]

50

例えば、上記説明した変形例を組み合わせてもよい。

【0211】

例えば、ユーザの不正度を予め取得できる場合には、不正度に応じて所持認証の方法を変えてもよい。不正度は、不正の度合いを示す情報、又は、不正の疑いの高さを示す情報である。ここでは、スコアによって不正度が表現される場合を説明するが、不正度は、他の指標で表現されてもよい。例えば、不正度は、Sランク・Aランク・Bランクといった文字で表現されてもよい。例えば、学習モデルMを利用して不正度が計算されてもよいし、ルールを利用して不正度が計算されてもよい。例えば、IPアドレスにばらつきがあるほど不正度が高くなるように、不正度が計算されてもよい。また例えば、ユーザがアクセスしたURLにばらつきがあるほど不正度が高くなるように、不正度が計算されてもよい。また例えば、アクセス場所が利用中心地から離れているほど、又は、アクセス場所にばらつきがあるほど、不正度が高くなるように、不正度が計算されてもよい。

10

【0212】

例えば、ユーザの不正度に基づいて、第1カードC2のICチップcpの記憶領域のうち、NFC認証で読み取る記憶領域が異なってもよい。例えば、ICチップcpは、読取部による読み取りのために鍵が必要な第1記憶領域と、読取部による読み取りのために鍵が不要な第2記憶領域と、を含む場合、ユーザの不正度が閾値以上であれば、第1記憶領域から入力電子マネーIDが取得されてもよい。ユーザの不正度が閾値未満であれば、第2記憶領域から入力電子マネーIDが取得されてもよい。この場合、第1記憶領域又は第2記憶領域の何れから入力電子マネーIDが取得されたかを示す情報が事業者サーバ30に送信され、所持認証において、この情報が確認されてもよい。

20

【0213】

また、ユーザの不正度に応じて、NFC部23A及び撮影部26の中から認証で利用するものが決定されてもよい。例えば、不正度が閾値以上の場合にNFC部23Aを利用すると決定され、不正度が閾値未満の場合に撮影部26を利用すると決定されてもよい。これとは逆に、不正度が閾値以上の場合に撮影部26を利用すると決定され、不正度が閾値未満の場合にNFC部23Aを利用すると決定されてもよい。他にも例えば、不正度が閾値以上の場合にNFC部23A及び撮影部26の両方を利用すると決定され、不正度が閾値未満の場合には、NFC部23A又は撮影部26の何れか一方を利用すると決定されてもよい。NFC部23A及び撮影部26のうち、認証で利用すると決定したものを識別する情報が事業者サーバ30に送信され、所持認証において、この情報が確認されてもよい。

30

【0214】

また、第1カードC2が複数の認証情報を含む場合に、ユーザの不正度に基づいて、認証で利用される認証情報が決定されてもよい。例えば、不正度が高いほど認証で利用する認証情報が多くなるように、認証で利用される認証情報が決定される。また例えば、不正度が低いほど認証で利用する認証情報が少なくなるように、認証で利用される認証情報が決定される。また例えば、不正度が閾値以上の場合に、比較的情報量の多い第1認証情報を利用すると決定され、不正度が閾値未満の場合に、比較的情報量が少ない第2認証情報を利用すると決定される。

【0215】

例えば、不正検知システムSは、行政サービス及び電子決済サービス以外の任意のサービスに適用可能である。例えば、不正検知システムSは、電子商取引サービス、旅行予約サービス、通信サービス、金融サービス、保険サービス、オークションサービス、又はSNSといった他のサービスにも適用可能である。第1実施形態の不正検知システムSを他のサービスに適用する場合、ユーザ端末20から所持認証等の所定の認証を実行した認証済みユーザの認証済み情報を利用して、学習モデルMが作成されるようにすればよい。第2実施形態の不正検知システムSを他のサービスに適用する場合も同様に、所持認証等の所定の認証を実行した認証済みユーザの認証済み情報を利用して、学習モデルMの精度を評価すればよい。

40

【0216】

50

例えば、所持認証で利用されるカードは、保険証、免許証、会員証、又は学生証等であってもよい。所持認証で利用されるカードは、物理的なカードではなく、電子的なカード（バーチャルなカード）であってもよい。また例えば、所持認証が失敗した場合には、管理者による人手の判定が行われてもよい。また例えば、あるカード番号に対応する所持認証が所定回数だけ失敗した場合には、そのカード番号については、それ以上の所持認証が実行されないように制限されてもよい。この場合、管理者による許可がない限りは、そのカードがアプリに登録されないように制限がかけられてもよい。他にも例えば、情報記憶媒体の読み取りによって所持認証が実行されてもよい。

【 0 2 1 7 】

例えば、主な機能がサーバ 1 0 又は事業者サーバ 3 0 で実現される場合を説明したが、各機能は、複数のコンピュータで分担されてもよい。

10

20

30

40

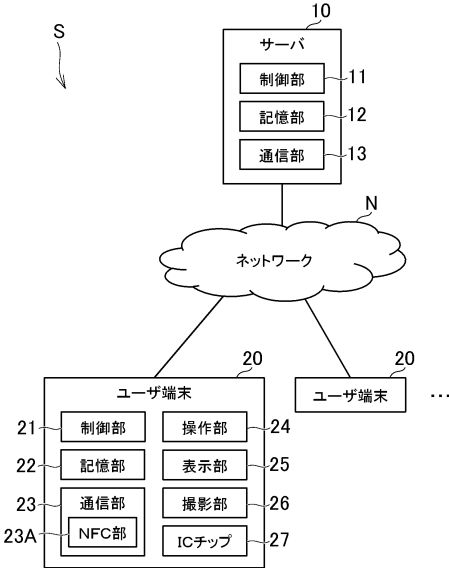
50

【要約】

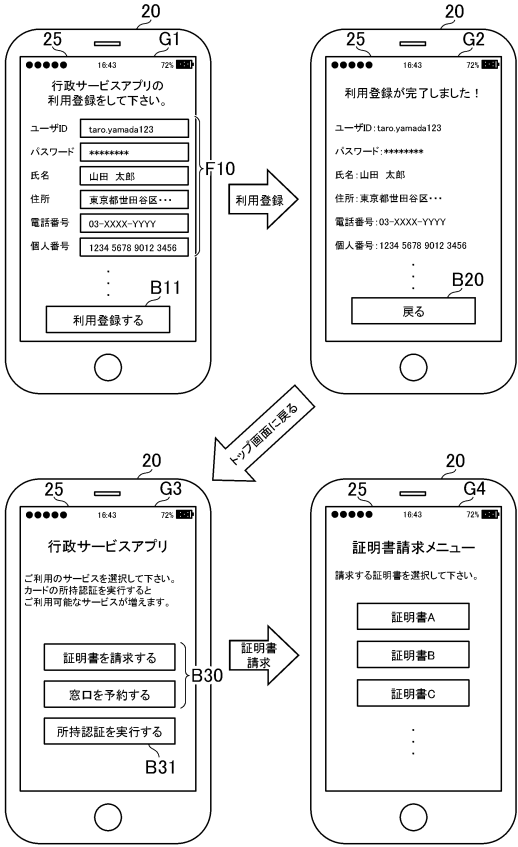
学習モデル評価システム（S）の認証済み情報取得手段（101）は、所定のサービスを利用可能なユーザ端末（20）から所定の認証を実行した認証済みユーザの行動に関する認証済み情報を取得する。出力取得手段（104）は、認証済み情報に基づいて、サービスにおける不正を検知するための学習モデルからの出力を取得する。評価手段（105）は、認証済み情報に対応する出力に基づいて、学習モデルの精度を評価する。

【図面】

【図 1】



【図 2】



10

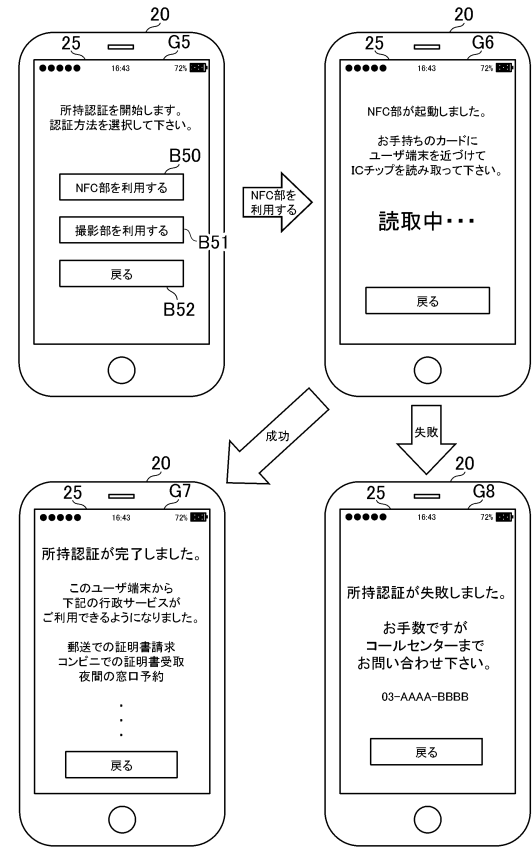
20

30

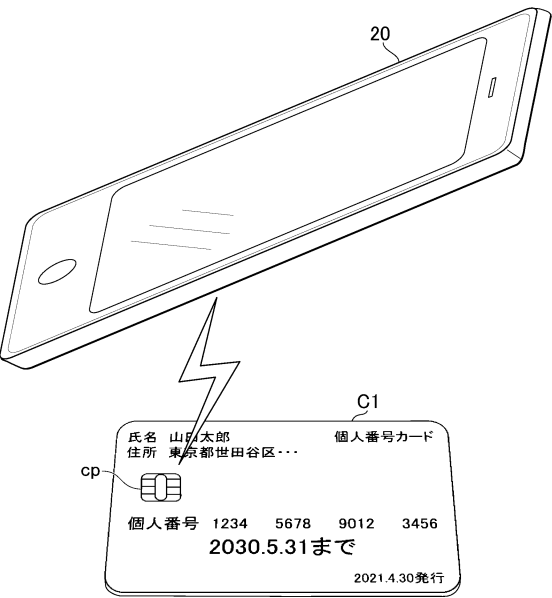
40

50

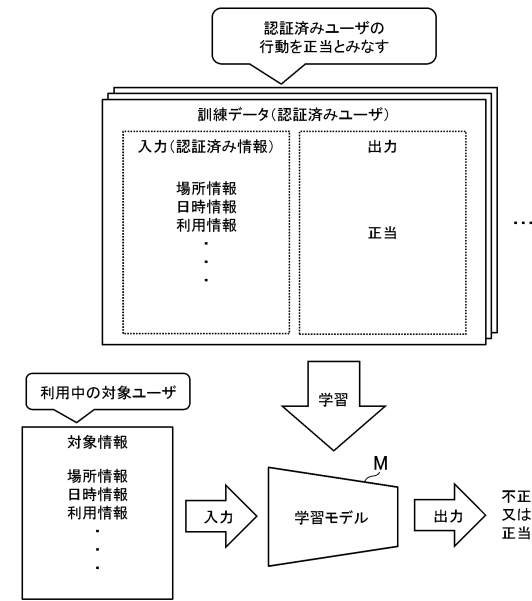
【図 3】



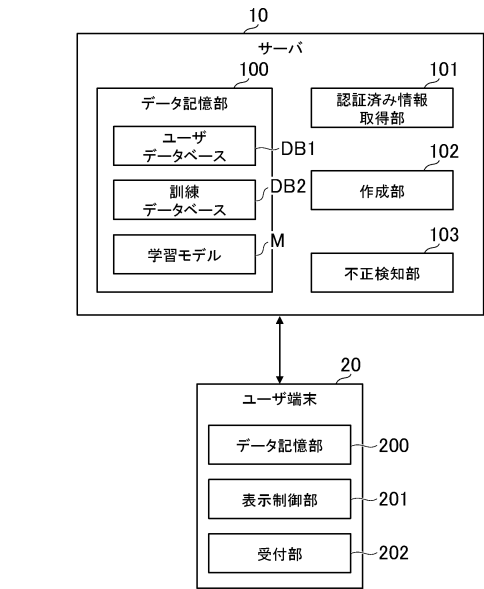
【図 4】



【図 5】



【図 6】



10

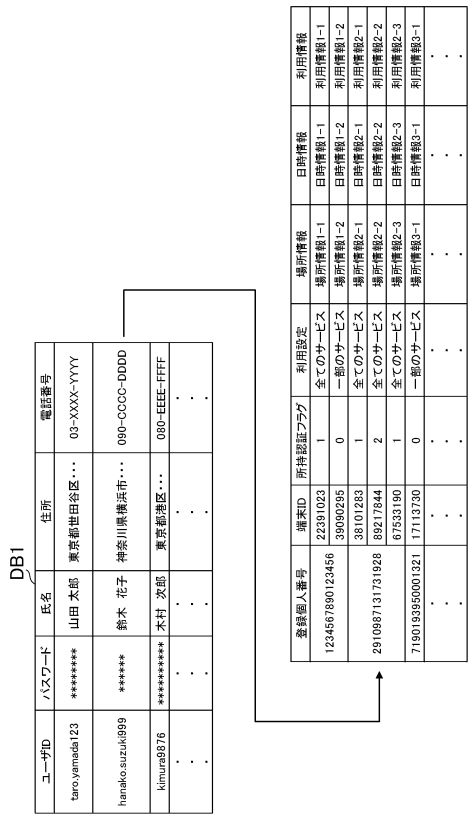
20

30

40

50

【図 7】



【図 8】

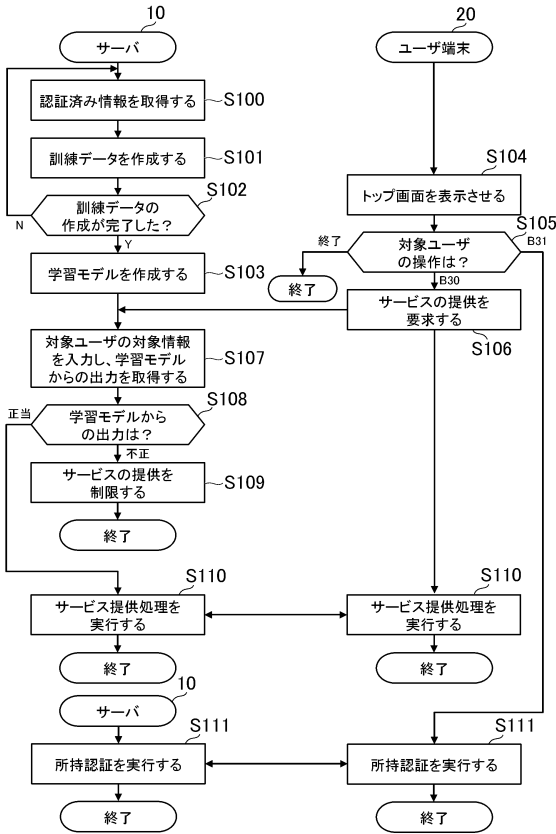
DB2

入力(認証済み情報)			出力
場所情報100	日時情報100	利用情報100	0
場所情報101	日時情報101	利用情報101	0
場所情報102	日時情報102	利用情報102	0
場所情報103	日時情報103	利用情報103	0
場所情報104	日時情報104	利用情報104	0
場所情報105	日時情報105	利用情報105	0
...	...	...	...
...	...	...	...

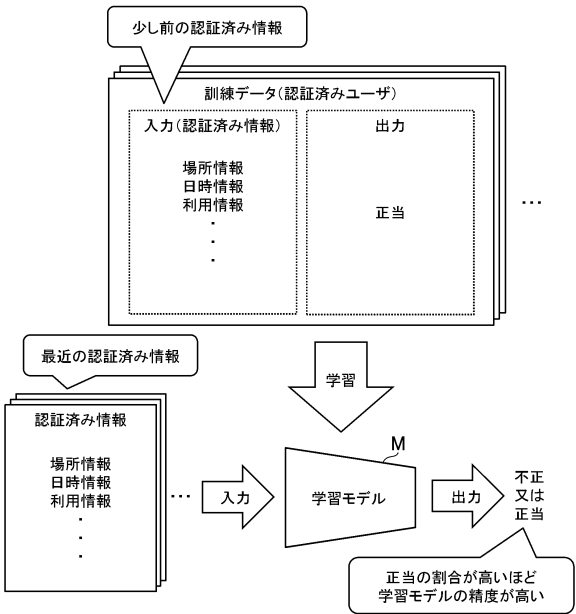
10

20

【図 9】



【図 10】

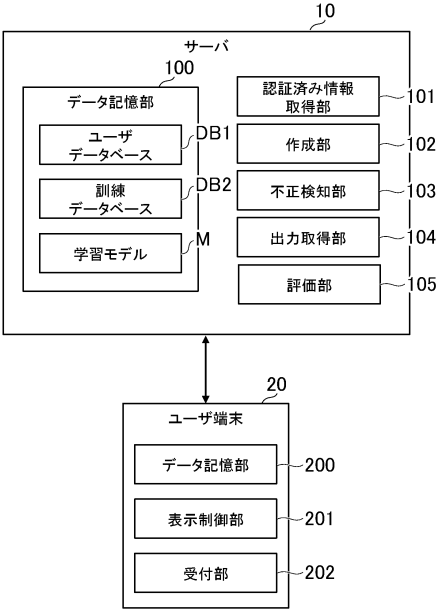


30

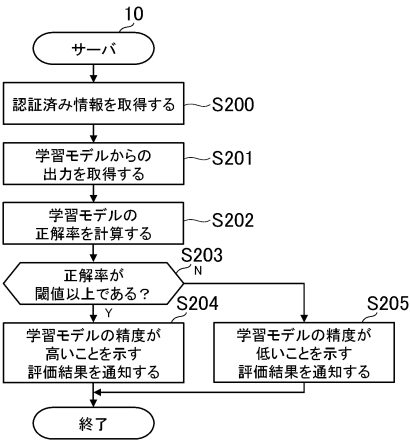
40

50

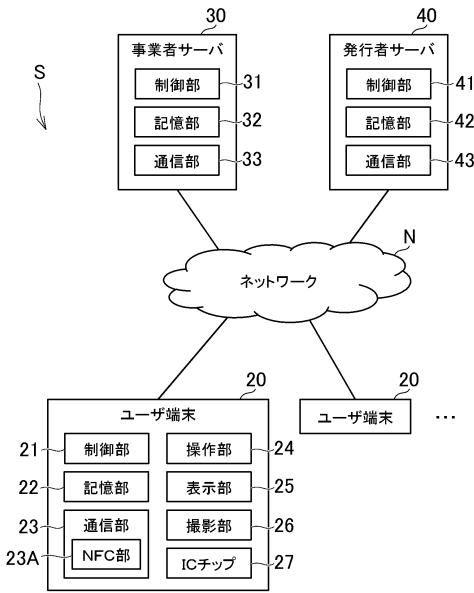
【図 1 1】



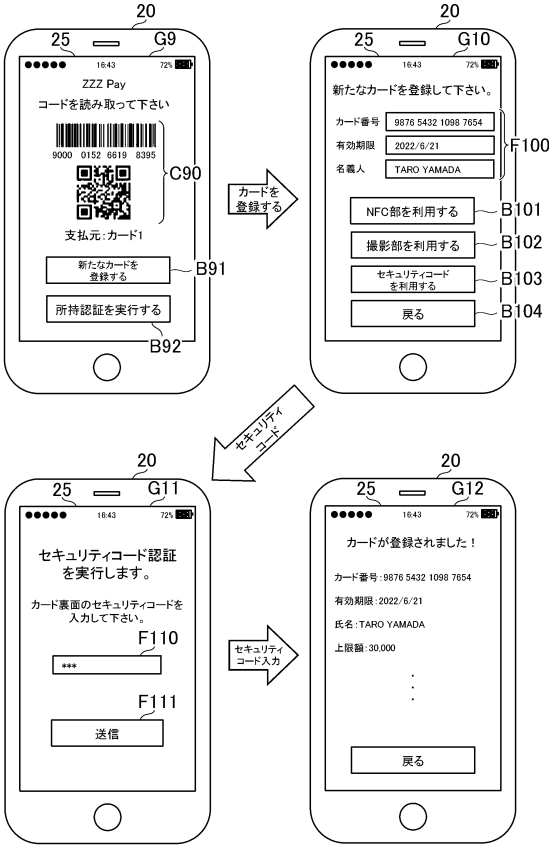
【図 1 2】



【図 1 3】



【図 1 4】



10

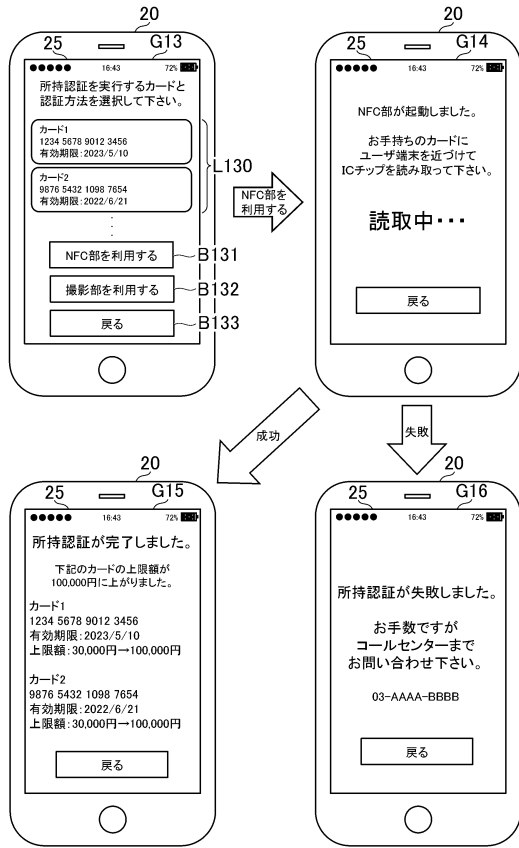
20

30

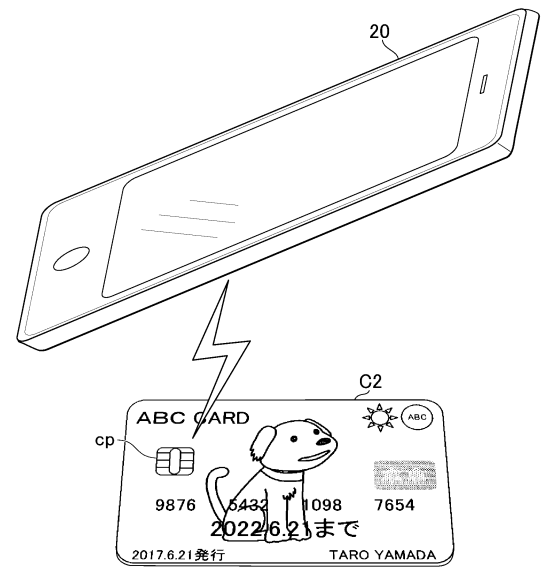
40

50

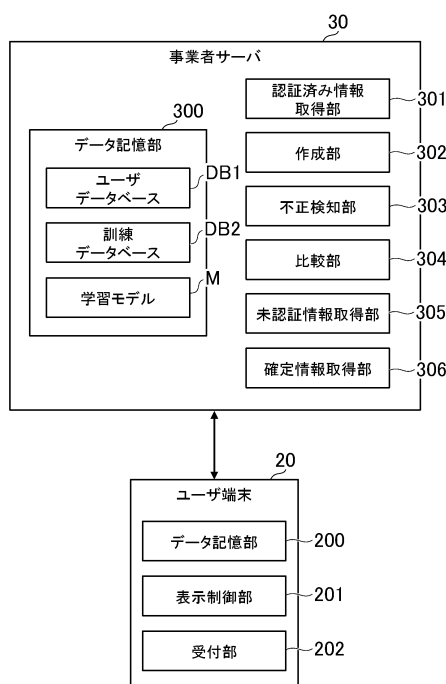
【図 15】



【図 16】



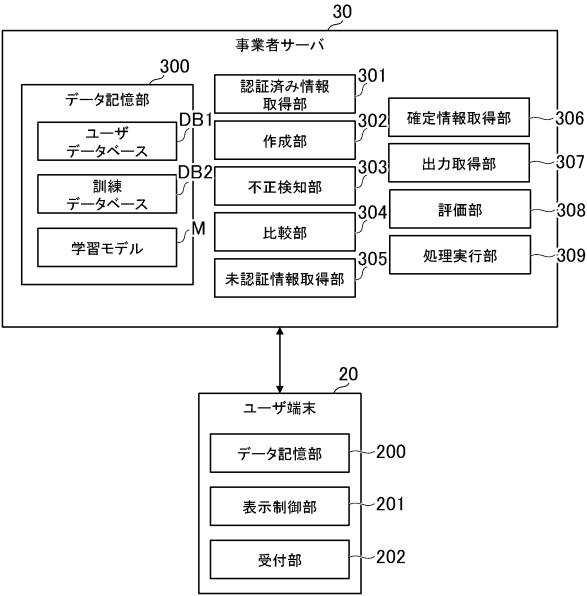
【図 17】



【図 18】

ユーザID	パスワード	氏名	支払元	登録カード情報				DB1				電子キャッシュ情報			
				No.	カード番号	有効期限	名義人	所持認証フラグ	利用限度	場所情報	日時情報	場所情報	日時情報	利用情報	利用情報
taro.yamada123	*****	山田 太郎	カード1	1	1234567890123456	2023/5/10	TARO YAMADA	0	10万円	場所情報1-1	日時情報1-1	場所情報1-1	日時情報1-1	利用情報1-1	利用情報1-1
hanako.suzuki099	*****	鈴木 花子	カード2	2	9876543210987654	2022/6/21	HANAKO SUZUKI	1	10万円	場所情報2-1	日時情報2-1	場所情報2-1	日時情報2-1	利用情報2-1	利用情報2-1
kimura0976	*****	木村 次郎	電子キャッシュ	3	7131317424324141	2023/2/15	MRI OKAMOTO	0	3万円	場所情報3-1	日時情報3-1	場所情報3-1	日時情報3-1	利用情報3-1	利用情報3-1

【図 19】



10

20

30

40

50

フロントページの続き

- (56)参考文献 特開 2019-008369 (JP, A)
国際公開第 2021/038775 (WO, A1)
特開 2014-167680 (JP, A)
特開 2020-115175 (JP, A)
特開 2005-258801 (JP, A)
石川 竜雄, マイナンバーカードを活用した多要素認証の実現, 月刊自動認識, 日本, 日本工業出版, 2019年12月10日, 第32巻 第14号, pp.24-26
- (58)調査した分野 (Int.Cl., DB名)
G06F 21/55
G06N 20/00