大学向けリスクベース認証アルゴリズムの検討

仲山 悠也[1]　笠原 禎也[1,2]　高田 良宏[2]　松平 拓也[2]　東 昭孝[2]

概要：金沢大学では，Shibboleth を用いた統合認証基盤を構築・運用している．現在は ID・パスワード認証だけの運用であるが，セキュリティ強化のために，リスクベース認証の導入を検討している．本研究では，本学において最適なリスクベース認証を実現するため，運用中の統合認証基盤のログデータの解析を行い，リスク判定基準と，大半の人に適用可能である汎用的なリスクベース認証アルゴリズムを提案し，シミュレーションによる評価・検証を行なった．

キーワード：セキュリティ，シングルサインオン，Shibboleth，リスクベース認証

Study on a Risk-Based Authentication Algorithm for Information System in University

Yuya Nakayama[1]　Yoshiya Kasahara[1,2]　Yoshihiro Takata[2]　Takuya Matsuhira[2]
Akitaka Higashi[2]

Abstract: We have been operating an integrated authentication infrastructure based on Shibboleth in Kanazawa University. Currently, we only use ID and Password for authentication, but we consider to introduce a risk-based authentication to enhance security. In this research, we analyzed the log data of the past authentication records in the authentication server, and propose some criteria and algorithms appropriate for the risk-based authentication. Finally, we summarize the evaluation results obtained by the simulation according to the proposed criteria and algorithms.

Keywords: Security, Single Sign-On, Shibboleth, Risk-Based Authentication

1. はじめに

近年，学生・教職員向けに提供する各種の全学情報サービスは，連携・融合化が進み，認証システムのシングルサインオン（以下，SSO とする）化が進んだ．その結果，認証に必要な ID とパスワードが一元化され，種々のサービスが同一の ID とパスワードで利用可能になった．反面，一度の認証で複数サービスにアクセス可能となることから，認証機構のセキュリティ強化が非常に重要視されている．認証時のセキュリティはあらゆる情報サービスにおいて重要であるため，ID・パスワード認証の他に，次世代認証方式である IC カードやスマートフォンと連携し，複数の要素を認証に使用する多要素認証や，通常とは異なる環境からの認証要求があった場合に追加の認証を要求するリスクベース認証，複数のパスワードを利用する多段階認証など様々な方法が提案されている．

金沢大学（以下，本学とする）においても，SSO システムとして，金沢大学統合認証基盤（Kanazawa University Single Sign-On（以下，KU-SSO とする））が，平成 22 年 3 月から本格稼働を開始している [1, 2, 3]．KU-SSO の構築により，システム管理者が各システムの認証機構やユーザを個別に管理する必要がなくなり，システム間での連携も取りやすくなった．現在は，ID・パスワード認証だけで運用しているが，サービスの連携拡大に伴い，よりセキュアな認証基盤の構築が必須である．

1　金沢大学大学院自然科学研究科電子情報科学専攻
Graduate School of Natural Science and Technology, Kanazawa University
2　金沢大学総合メディア基盤センター
Information Media Center, Kanazawa University

本研究では，大学の認証基盤をより強固なものとするため，次世代認証方式としてリスクベース認証の導入へ向けて検討を行う．そのために，過去のログデータを解析し，本学のユーザの利用形態に適したリスクベース認証を実現するための判定基準と，アルゴリズムの検討結果について報告する．そして，提案するリスクベース認証アルゴリズムについて，過去のログデータを利用したシミュレーションを行い，その妥当性の評価結果についてまとめる．

2. 金沢大学統合認証基盤

本章では，金沢大学が構築・運用している統合認証基盤である KU-SSO について解説する．

2.1 概要

本学の統合認証基盤である KU-SSO は，Shibboleth[4] による認証・認可機能を採用し，一度の認証で，そのユーザが利用する権利を持つ全ての学内サービスが利用できる SSO を実現している．Shibboleth は主に Identity Provider（以下，IdP とする），Service Provider（以下，SP とする），Discovery Service（以下，DS とする）の 3 つのサーバから構成される．IdP は大学などの組織単位で構成され，ユーザの認証を行うサーバであり，SP は IdP からの情報を元にユーザに対して各種サービスを提供するサーバである．DS は IdP のリストを保持し，ユーザに自身の所属する組織の IdP を選択させる機能を持ったサーバである．本学のユーザであれば，本学が構築・運用する IdP によってユーザ認証を行い，各種 SP は IdP から受信したユーザの属性情報を元にサービスを提供する．

本学のユーザは，金沢大学 ID とパスワードで認証を行う．金沢大学 ID とは，学生・常勤教職員・非常勤教職員などを問わず，本学に関わる構成員全てに付与される生涯 ID であり，3 つのアルファベットと 5 つの数字の 8 文字で構成される．現在，本学の各 SP への入り口として機能する「アカンサスポータル」をはじめとした，30 以上の SP の認証を請け負っており，学生・教員の連絡手段となる「お知らせ」や，「メッセージ」など様々なサービスを提供している [5]．また，本学は学術認証フェデレーション [6]（以下，学認とする）に参加し，学認に参加している外部機関のシステムとも認証連携を行うことが可能である．日々のログイン回数は，平日約 14,000 回，休日は約 2,500 回であり，履修登録期間などは最大で 1 日に約 20,000 回のアクセスがある．

2.2 KU-SSO の現状

本学内で提供される SP 群は，学内からは全てアクセス可能だが，給与明細 SP や予算執行支援 SP などの重要度の高い SP は，学外 IP アドレスからのアクセスを制限している．理由は，提供されている認証方式が ID・パスワード認証のみであることによるセキュリティ上の懸念があるからである [7]．学外からこれらの SP を利用したい場合は，金沢大学 ID とは異なる ID（ネットワーク ID）を用いた認証を必要とする VPN を使用し，学内ネットワークに接続してから SP にアクセスする必要がある．しかし，VPN との併用は，一般ユーザにとってわかりにくく利便性に欠ける問題や，SP の認証強化とは本来無関係である VPN を認証強化に代用していることに問題がある．SP を外部に公開し，利便性を向上させるためには，認証機構によって適切な認証レベルを設定可能とすることが必要である．

この問題に対処するため，本学では多要素認証の開発が進められてきた [8]．多要素認証とは，本人しか知らない知識（パスワード，PIN など），本人しか持っていない所有物（IC カード，スマートフォンなど），本人の生体的特徴（指紋，静脈など）の 3 つの要素のうち，2 要素以上を必要とする認証方式である．例えば，IC カード（所有物）と PIN（知識）といった組み合わせで認証が行われる．このように，認証時に 2 要素以上の組み合わせが必要なため，セキュリティ強度の向上が可能である．本学における多要素認証には，スマートフォン（所有物）と PIN（知識）による tiqr 認証 [9] 及び，YubiKey デバイス [10]（所有物）と ID・パスワード（知識） による YubiKey 認証の導入が検討されている．

2.3 大学向け次世代認証方式の検討

多要素認証では，より高い認証レベルの設定が可能となる反面，認証時の手間は ID・パスワード認証よりも増加する．そのため，大学のシステムを構成する大規模な SP 群に一括して多要素認証を導入することは，ユーザの利便性の低下を招くことに繋がり，現実的ではない．他大学においては，アクセス元が学外の場合に多要素認証を要求するなど，利便性と安全性のバランスを保つ工夫を施している [11]．

一般的に多要素認証では，各 SP 毎にあらかじめ必要とする認証レベルを定める必要がある．つまり，重要な SP には高い認証レベルを設定し，一般的な SP では単一要素のみ要求するといった具合である．しかし，より良い認証基盤を実現するためには，真に必要な場面でのみ多要素認証を求めるなど，より柔軟に動作可能な認証方式を検討することが必要である．

このような背景から，本研究では，大学向けの次世代認証方式として，ユーザの行動履歴を利用し，動的に多要素認証を要求可能な認証システムの構築へ向けて検討を行なった．

3. 大学向け次世代認証方式

本章では，大学向け次世代認証方式として着目したリスクベース認証の概要と，利用するリスク判定基準について

検討する．

3.1 リスクベース認証

認証時にユーザの環境情報や行動パターンを分析し，リスクを判定した上で，必要であれば追加認証を課すような認証方式は，リスクベース認証と呼ばれている．具体的な例として，SSO を実現するソフトウェアである ForgeRock Access Management（元 OpenAM）[12] では，リスクベース認証を実現するモジュールとして以下のものが提供されている．

- Adaptive Risk Authentication
  - 最終ログインからの経過時間や，IP アドレスの履歴，認証の失敗回数などから本人ではないリスクを判定する認証モジュール．
- Device ID Authentication
  - ブラウザなどのユーザーエージェント，インストールされているフォント，画面解像度や色深度などのデバイス特性から本人ではないリスクを判定する認証モジュール．

リスクベース認証は，普段とは異なるユーザの行動を検知した場合に追加認証を要求する．不正アクセスは，ユーザが普段活動している場所や，普段利用するデバイス以外からのアクセスなど，普段とは異なる行動がある可能性が高いため，それらを検知して追加認証を要求するリスクベース認証は，安全性の向上が期待できる．

3.2 リスク判定基準

本研究では，ID・パスワード認証をベースにリスクベース認証を導入し，リスクベース認証の追加認証に多要素認証を用いることによって，安全性と利便性を兼ね備えた認証基盤を構築することを目標とする．この方法では，普段の利用では ID・パスワード認証のみ通過すれば良く，ユーザの利便性を損なわない．それでも，何か異常を検知した際には，追加で多要素認証が要求されるため，セキュリティ強度の向上が期待できる．

KU-SSO は，ユーザの立場によってサービスの使い方や使用頻度が多種多様である．そのため，様々な利用形態のある KU-SSO のユーザにリスクベース認証を適用した際に，誰かが著しく利便性を損なうことがないように，リスクの判定基準を慎重に定める必要がある．しかし，当然異常を異常として検知できなければ意味がない．

そこで本研究では，本学のユーザがサービスをどのような環境で利用しているか調査を行い，利用実態に適したリスク判定基準を定めるために，過去の IdP のログデータを利用したデータ解析を行なった．

4. ログデータ解析

本章では，IdP のログデータを利用した解析と，その結果について考察する．

4.1 利用情報

解析には，IdP の一連の処理が記録されているプロセスログを使用し，主に 2014 年度と 2015 年度の 2 年間分のログデータを対象とした．ユーザのアクセス時に関する情報から，リスクベース認証に利用可能と考えられる以下 4 つの属性情報を抽出した．

- requestTime（アクセス日時）
- remoteHost（アクセス元 IP アドレス）
- relyingPartyId（アクセス先 SP）
- principalName（金沢大学 ID）

本研究では，これらの情報の中から，特に IP アドレスの情報を活用し，ユーザの日常利用と非日常利用を判別するための検討を行なった．

4.2 ネットワークデータベースの作成

ユーザが利用するネットワークの IP アドレス情報をそのまま利用してリスクを判定することは，過去にアクセスのあった IP アドレスを全て記録することが必須となる上に，DHCP 環境などでネットワークを利用するユーザは，使用する IP アドレスが常に同一とは限らないため，現実的ではない．そこで，個人のアクセス履歴に記録されている IP アドレスを，事業者（以下，ネットワークとする）ごとに分類した．個々の IP アドレスが属するネットワークが特定できれば，ユーザが日常的に利用するネットワークからのアクセスと，そうではないネットワークからのアクセスを区別可能となる．

IP アドレスをネットワークに分類するため，ログファイルに記録された 2014 年度分の全 IP アドレスについて whois による問い合わせを行い，IdP へのアクセス実績があるネットワーク名に関するデータベースを作成した．ログファイルには 3,569,904 件のアクセスログが記録されており，アクセスがあった IP アドレスは 145,088 件であった．これらの IP アドレスを JPNIC[13] の WHOIS サーバに問い合わせを行い，ネットワーク情報を取得した．そのため，JPNIC からネットワーク情報を得られなかった場合は，基本的にはその IP アドレスのネットワーク情報は記録されていない．現時点（2017 年 9 月）では，688 件のネットワーク情報がデータベースに記録されている．データベースに存在しない IP アドレスからのアクセス数は 7,687 件であり，割合は 0.22 %であった．作成したデータベースには，ネットワーク名と，CIDR 表記によるネットワーク情報を格納した．例えば，学内ネットワークの場合，ネットワーク名が「Kanazawa University」，ネットワーク情報は「133.28.0.0/16」と記録した．このデータベースを作成したことにより，IP アドレスからネットワーク名を得ることが可能となった．

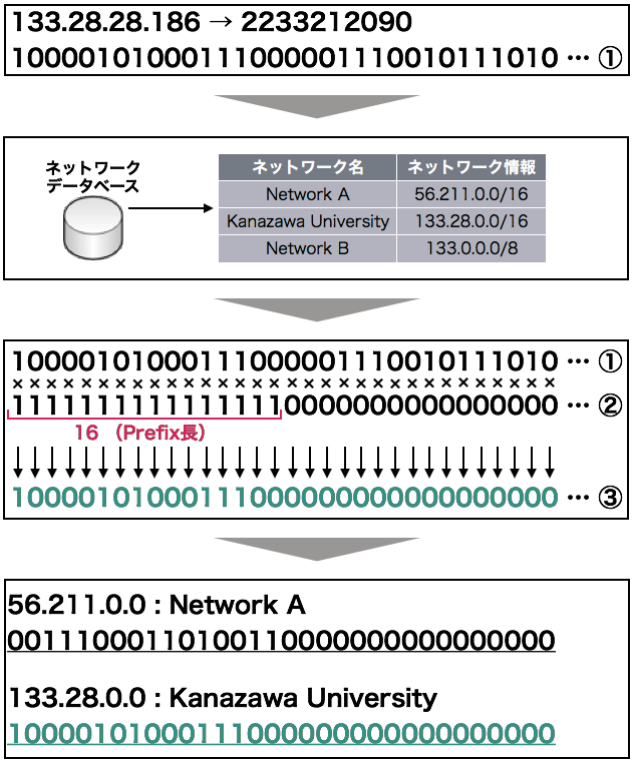

図 1　ネットワーク名取得までの流れ
Fig. 1　Acquisition process of a network name.

4.3　データベースを用いたネットワーク名取得手順

作成したネットワークデータベースを用いて，IP アドレスからネットワーク名を得るまでの手順を以下に示す．また，具体例を図 1 に示す．図 1 は，“133.28.28.186”という IP アドレスのネットワーク名を取得する際の流れを簡略化して表したものである．ただし，図 1 に登場するネットワークは説明目的のダミーである．

(1) IP アドレスを 32 ビット整数に変換する（①）．
(2) ネットワークデータベースからネットワーク名とネットワーク情報を取得する．
(3) ネットワークが小さいものから順番に以下の操作を繰り返す．
  (a) ネットワーク情報からプレフィックス長 P を読み取る．
  (b) 上位 P ビットが 1 の 32 ビット整数（②）と①の論理積をとる．
  (c) 取得したネットワークを 32 ビット整数に変換したものと，③を比較し，等しければネットワーク名を得て終了し，異なれば次のネットワークで（a）へ戻る．
(4) データベースのネットワークを全て走査し，該当するネットワークがなければ「該当なし」として終了する．

図 1 では，③の結果が，“133.28.0.0”を 32 ビット整数化したものと等しくなるため，ネットワーク名として“Kanazawa University”を得る．ネットワークを小さな順番に比較する理由は，大きなネットワークが小さなネットワークを内包する場合があるためであり（図 1 中の“Network B”は“Kanazawa University”を内包している），その場合に，より小さなネットワークを分類結果とするためである．このデータベースを利用し，個人のアクセス履歴の IP アドレスネットワーク名と紐付けて集計を行うと，ネットワーク別のアクセス数を得ることができる．

表 1　ユーザが利用するネットワーク別アクセス数の代表例（1）
Table 1　An example of network names and access frequency of a typical user (1).

Network	Count
Kanazawa University	206
ネットワーク A	58
ネットワーク B	19
ネットワーク C	9
合計	292

表 2　ユーザが利用するネットワーク別アクセス数の代表例（2）
Table 2　An example of network names and access frequency of a typical user (2).

Network	Count
ネットワーク D	126
Kanazawa University	54
ネットワーク E	45
ネットワーク F	1
合計	226

4.4　ネットワークに基づくユーザ分析

作成したデータベースを用いて，ユーザのネットワークに基づく分析を行った．対象は，2015 年度の 1 年間のログデータであり，ユーザ数は 15,435 人である．このうち，利用ネットワークが 1 つのユーザは 19 %，2 つのユーザは 18 %，3 つのユーザは 24 %，4 つのユーザは 21 %，5 つのユーザは 12 %，それ以上のユーザは 7 %であった．ユーザごとのネットワーク別アクセス数の例として，代表的なユーザの結果を表 1，2 に示す．これらは，全ユーザのアクセス数の平均値である 237 回付近のユーザの結果である．ユーザの全体的な傾向としては，表 1 のように，“Kanazawa University”からのアクセス数が最も多いユーザが多く，ユーザ全体の 54 %を占めている．また，そうではないユーザについては，表 2 のように，最も頻繁に利用するネットワークのアクセス数が，学内ネットワークのアクセス数に大きく差をつけている例が多く見られた．

ユーザの調査を行うと，大半のユーザにおいて，ネットワークの利用頻度に偏りがあり，普段からよく利用するネットワークとそうではないネットワークが存在していた．これを利用し，ユーザの日常利用とそれ以外を区別するために，アクセス数の上位にあるネットワークが，そのユー

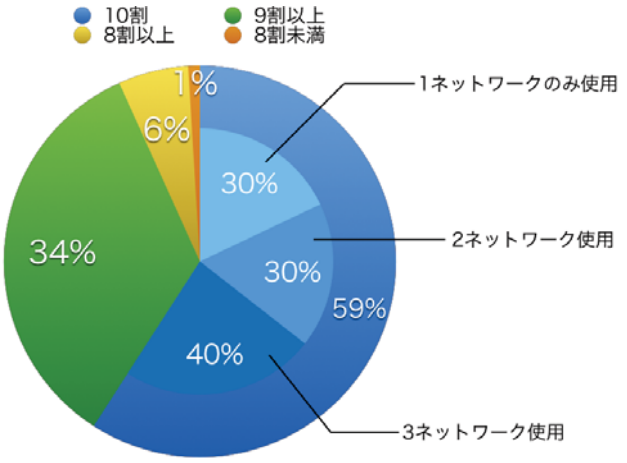

図 2　上位 3 つのネットワークを使用しているユーザの割合
Fig. 2　Percentage of users using the top 3 networks.

ザの全アクセス数の何割を占めるかについて調査した．その結果を図 2 に示す．

図 2 の青色で示した部分は，利用しているネットワークが 3 つ以下のユーザの割合である．そのようなユーザは全体の 59 %で，内訳は，利用ネットワークが 1 つであるユーザと 2 つであるユーザはそれぞれ 30 %，利用ネットワークが 3 つのユーザは 40 %であった．緑色，黄色，橙色で示した部分は，利用しているネットワークが 4 つ以上のユーザの割合である．緑色で示した部分は，上位 3 つのネットワークからのアクセス数が，全アクセス数の 9 割以上 10 割未満に該当するユーザである．同様に，黄色部分は 8 割以上 9 割未満，橙色部分は 8 割未満のユーザの割合である．

図 2 の結果から，上位 3 つのネットワークからのアクセスが各個人のアクセス数の 9 割以上であるユーザは全体の 93 %に及ぶことが確認できる．これは本学の情報サービスを利用するユーザの利用実態が，大学内からの利用と，自宅からの利用と，モバイルデバイスからの利用の 3 つのネットワーク環境がほとんどであるためと推察できる．そのため，大半のユーザにおいて，日常的に利用するネットワークの数は高々 3 つであり，上位 3 つに入らない残りのネットワークは日常的に利用するネットワークではないとみなすことが可能であると考えられる．図には掲載していないが，同様の分析を上位 2 つのネットワークで行うと，9 割以上のユーザは 65 %，上位 4 つのネットワークで行った場合は 99 %となった．

この結果に基づき，以下のアルゴリズムの検討では，個々のユーザが利用する日常的なネットワークの数は，高々 3 つであるという前提で検討を進めることとした．

5.　リスクベース認証アルゴリズムの検討

本章では，ユーザの大学情報サービスの利用実態に基づいて最適化したリスクベース認証のアルゴリズムについて検討と評価を行う．

5.1　認証ポリシー

本研究では，以下のようなアクセスに対して，追加認証を要求すべきであるとする認証ポリシーを作成した．

- ネットワークデータベースに登録がないネットワークからのアクセス
  - whois でネットワークが特定できなかった 0.22 %のネットワークからのアクセスがこれに該当する．
- ユーザが初めて利用するネットワークからのアクセス
- 最終アクセスから一定期間以上のアクセス
- ユーザが日常的に利用するネットワーク以外からのアクセス
  - 判断基準には，4 章で示したネットワーク利用実態に基づくリスク判定を利用する．

この認証ポリシーの元，リスクベース認証のアルゴリズムを検討する．

5.2　リスクベース認証アルゴリズム

認証ポリシーに基づき，検討したリスクベース認証アルゴリズムの流れを図 3 に示す．アルゴリズムによって追加認証と判定された場合には，多要素認証による追加認証をユーザに求める．検討したアルゴリズムにおいて，最初の①～③は全ユーザに対して適用される内容であり，最終アクセスからの経過時間の検証と，IP アドレスのネットワークが既知かどうかの検証，これまでのアクセス件数の検証が行われる．①，②において要件を満たさない場合には，追加認証がユーザに課される．③においてアクセス数が一定基準を満たした場合は，一定回数以上の利用があるユーザに対して行うリスクベース認証として④が実行され，アクセス元ネットワークの情報から日常利用かどうかを検証する．これにより，日常利用のネットワークと，そうではないリスクの高いネットワークの切り分けを行う．しかし，非日常的なアクセスであっても，出張先からのアクセスのように，特定の期間に限って，ある特定のネットワークから複数回のアクセスをするケースがありうる．そのような場合に，毎回リスクベース認証を行うことは，ユーザに過度な負担を強いることになる．そこで，⑤によって 1 度目のアクセスに限ってリスクベース認証を実施し，その後，一定期間は同ネットワークからのアクセスは通常認証のみで可とするアルゴリズムとした．アクセス数が一定基準に満たない場合は，日常利用のネットワークについて情報が不足しているため，⑥において対象ユーザにとってアクセス元のネットワークを利用したアクセスが初回であるかのみを検証するアルゴリズムとした．また，学内アクセスは無条件に許可するなどのルール設定も可能だが，今回は簡単のため，そのような想定は含まない形で検討を行なった．

5.3　アルゴリズムのシミュレーションによる検証実験

検討したリスクベース認証のアルゴリズムを導入した際

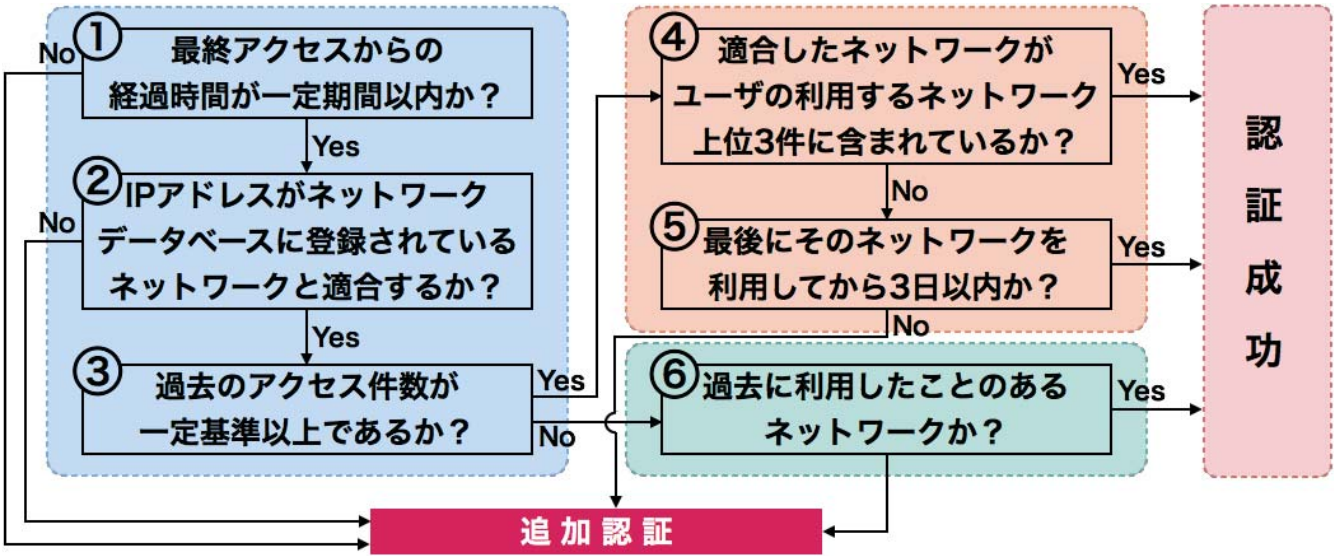

図 3　リスクベース認証アルゴリズム
Fig. 3　Risk-Based Authentication Algorithm

に，どのようなユーザやアクセスに対して影響を及ぼすのか，またその結果が妥当であるかを検証するため，実際のログデータに対してアルゴリズムを適用し，シミュレーションを行なった．シミュレーションを行うにあたり，いくつかのパラメータを設定した．最終アクセスからの経過時間の許容期間は，セキュリティポリシーの観点から約 1 ヶ月間の 30 日とした．また，過去のアクセス件数の基準値には，ユーザの 1 年間のアクセス数の平均値の 10 %程度である 20 件と設定した．シミュレーションは 2014 年度と 2015 年度の 2 年間のログデータを使用し，対象期間中にアクセスのあったユーザの中からランダムに抽出した 1,000 人のユーザを対象とした．Python によるプログラムを用いて，ユーザの 2 年間のアクセスログを古いものから順番にリスクベース認証のアルゴリズムに適用することでシミュレーションを行なった．

シミュレーション結果について，ユーザの分布をプロットしたものを図 4 に，また結果の概要をまとめたものを表 3 に示す．図 4 は横軸にアクセス回数，縦軸にリスクベース認証によって追加認証を要求された回数をプロットしたものである．表 3 は，ユーザのアクセス数，追加認証要求回数，使用しているネットワークの数について，それぞれ平均値・標準偏差・最小値・パーセンタイル値・最大値を求めたものである．パーセンタイル値とは，ユーザを各項目ごとに昇順でソートした際の下から 25 %，50 %，75 %の位置にいるユーザの値を取り出したものである．今回のシミュレーションの対象者は 1,000 人であるため，アクセス数，追加認証要求回数，ネットワーク数の各項目においてユーザを昇順にソートし，下から 250 番目，500 番目，750 番目の人の値を取り出している．

アクセス数について，平均的には 2 年間でおよそ 465 回のアクセスがあることが読み取れる．また，標準偏差が約 623 と人によって非常にばらつきが多く，最大では 5,127 回のアクセスがあることがわかる．さらに，50 %点の値が約 284 であるため，半分以上を占める多くのユーザはアクセス数が平均に満たない位置に分布していることが表 3 と図 4 から読み取れる．

表 3　シミュレーション結果概要
Table 3　An overview of simulation results.

	アクセス数	追加認証要求回数	ネットワーク数
平均値	465.15	11.21	4.06
標準偏差	623.22	12.25	2.32
最小値	1.00	1.00	1.00
25 %点	100.75	5.00	3.00
50 %点	283.50	8.00	4.00
75 %点	538.00	13.00	5.00
最大値	5,127.00	206.00	36.00

追加認証要求回数については，平均的には 2 年間で約 11 回の追加認証が要求されていることがわかる．また，傾向としてはアクセス数と同様に，50 %点の値が 8 であることから，半分以上のユーザは平均以下に分布していることがわかる．最大 206 回の追加認証要求を受けているユーザなど，一部突出しているユーザについては，次節にて考察を行う．

ネットワーク数については，ユーザは平均 4 つのネットワークを利用していることがわかる．また，パーセンタイル値からも大半のユーザについて利用するネットワークの数は 3 つから 5 つであることが読み取れる．

5.4　考察

アクセス数のばらつきは，ユーザが学生や教職員など様々な立場の人が存在することが原因と考えられる．特

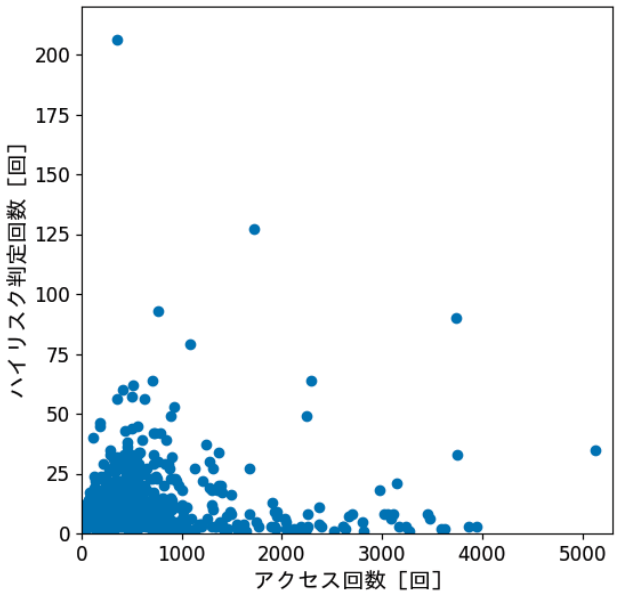

図 4　ユーザの分布
Fig. 4　A distribution of users.

に，アクセス数が数千回に及ぶユーザは，様々なサービスを頻繁に利用する教職員である可能性が高い．追加認証を要求されているアクセスについて調査すると，次のような特徴を持つ人が追加認証を要求される回数が多いことを確認した．

- データベースに存在しないネットワークをよく利用するユーザ
- あまり頻繁にサービスを利用しないユーザ
- 使用するネットワークの数が多いユーザ

現在のアルゴリズムでは，データベースに存在しないネットワークからのアクセスに対して追加認証を要求することにしているため、そのようなネットワークを主に使用しているユーザは追加認証を要求される回数が極端に増加している．図 4 において 200 回以上追加認証を要求されるユーザなど，突出しているユーザはこれに該当する場合が多い．あまり頻繁にサービスを利用しないユーザは，最終アクセスからの経過時間が 30 日以上開くことによって追加認証を求められている．アクセス数はそれほど多くないが，多く追加認証を要求されているユーザの場合はこれに該当することが多い．また、単純に使用するネットワークが多いユーザは，ネットワークに基づくリスク判定によって追加認証を要求される場合が増加する．このタイプのユーザは，出張先からのアクセスが多いユーザや，5 つ程度のネットワークを日常的に利用するユーザなどが該当する．前者であればリスクベース認証としては正しい動作である場合が多いが，後者の場合は必ずしもそうとは言い切れない．特殊なケースでは，ネットワークの使用頻度において上から 3 番目と 4 番目が同程度であった場合，両者が頻繁に入れ替わることによって追加認証を求められている場合などがあった．

実際の一ユーザのアクセス分析結果を図 5 に示す．このユーザは，2 年間で 155 回のアクセスがあり，そのうちリスクベース認証アルゴリズムによって 10 回の追加認証を求められているユーザである．また，このユーザが利用しているネットワークは 4 つであった．図 5 は，横軸に日付，縦軸にアクセス回数として，30 日間のアクセス数の推移を青色で示し，リスクベース認証アルゴリズムによって追加認証を課されたアクセスがあった日付を水色で示している．30 日間のアクセス数とは，図 5 の左端の値を例に挙げると，2014 年 5 月 1 日から，30 日間遡った日付までにあったアクセスの合計を示している．つまり，青色で示されている値が 0 になっている部分は，30 日間サービスの利用がなかったことを示している．シミュレーションでは，最終アクセスからの経過時間が 30 日以上の場合には追加認証を要求するため，青色で示した値が 0 から 0 以上になる部分には必ず水色の線が引かれており，追加認証が要求されていることが読み取れる．

図 5 では，9 月から 10 月にかけてアクセスがしばらくない時期があることがわかる．これは，この期間が夏季休暇に該当し，サービスを利用する必要がなかったからだと考えられる．一般に，多くのユーザにおいて，長期の休暇がある時期（8・9 月，12・1 月，3 月など）には，30 日間のアクセス数が 0 になる場合が多く見受けられる．また，これらの期間では，アクセス数が 0 にならない場合であっても，リスクベース認証による追加認証が要求されている場合が多い．これは，長期休暇中の帰省などによって，普段あまり利用しないネットワークからアクセスする機会が増えるためであると推察できる．

これらの分析結果から，利用するネットワークの数が 4 つである平均的な学生ユーザであれば，各ネットワークからの初回アクセスによって 4 回，夏季休暇などの長期休暇後のアクセスによって年 2 回程度，帰省先などから数回アクセスがあると考えれば，平均 11 回という追加認証を要求される回数は妥当であると言える．しかし，年間数千回のアクセスのある教員や，現在ネットワークデータベースに登録がないネットワークからの利用がメインのユーザなど，一部ユーザに対してはうまく機能しない場合もあるため，これらのユーザに対しては特別措置による対応が必要であると考える．

5.5　実装に向けた検討課題

リスクベース認証のアルゴリズムについて大筋がまとまったため，これからは既存の認証システムを構築している Shibboleth との連携を考える必要がある．Shibboleth IdP は Java で記述されており，Spring[14] というフレームワーク上に構成されている．特に，一連の手続き的な処理の実装には，Spring Web Flow[15]（以下，SWF とする）という Spring の機能が使用されている．そのため，Shibboleth

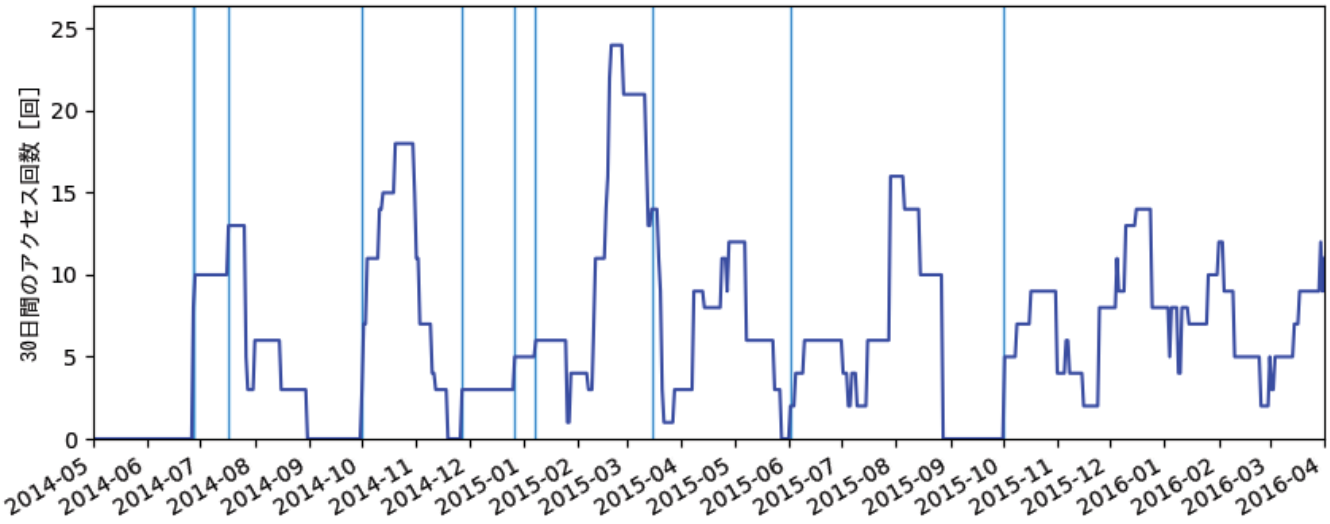

図 5　アクセス数の推移
Fig. 5　Changes in the access count.

にリスクベース認証を導入するには，リスクベース認証のための SWF を記述し，それが適切に実行されるように設定及び実装する必要がある．

また，実稼働の際には，ネットワークデータベースの運用方法についても検討する必要がある．例えば，未知のネットワークからのアクセスがあった場合，データベースにレコードを追加してもアルゴリズムの本質的な動作には影響しないため，未知の IP アドレスを whois で問い合わせ，結果が得られた場合はデータベースに随時追加する予定である．さらに，実稼働においてより効率化を求められる場合は，トライ構造を用いて効率的にネットワークを探索できるようにするなどの工夫が必要である．

6. まとめ

本研究では，リスクベース認証と多要素認証を組み合わせ，ユーザの利便性の低下を抑えて安全性を向上させる認証方式について検討を行なった．その上で，本学において最適なリスクベース認証を実現するため，IdP の過去のログデータの解析を行い，ネットワークに基づいたリスク判定基準と，大半の人に適用可能である汎用的なリスクベース認証アルゴリズムを提案し，シミュレーションによる評価・検証を行なった．シミュレーションの結果，しばらくサービスの利用がないユーザの不意のアクセスや，利用したことのないネットワークからのアクセスなどの不審な場面において，リスクベース認証のアルゴリズムが機能していることを確認し，そのようなアクセスに対して多要素認証を求めることにより，安全性の向上の実現に繋がることを示した．

今後は，KU-SSO にリスクベース認証を導入するため，Shibboleth 認証との連携方法や，設計及び実装について検討する予定である．

参考文献

[1] 松平 拓也，笠原 禎也，高田 良宏，東 昭孝，二木 恵：“学認との融合化を視野に入れた金沢大学統合認証基盤の構築と運用”，学術情報処理研究，No.16, pp.41-50, 2012.9
[2] 松平 拓也，笠原 禎也，高田 良宏，東 昭孝，二木 恵，森 祥寛：“大学における Shibboleth を利用した統合認証基盤の構築”，情報処理学会論文誌，Vol.52 No.2, pp.703-713, 2011.2
[3] 松平 拓也，笠原 禎也，高田 良宏，東 昭孝，二木 恵，藤田 翔也：“金沢大学における統合認証基盤の現状と課題”，大学 ICT 推進協議会 2013 年度年次大会（AXIES2013）論文集，W3E-4(CD-ROM), 18-20 December, 2013.
[4] Shibboleth: https://shibboleth.net (accessed 2017.09)
[5] 二木 恵，東 昭孝，笠原 禎也，高田 良宏，松平 拓也：“全学ポータルを用いた学生・教職員間多機能連絡システムの開発”，学術情報処理研究，No.16, pp.15-24, 2012.9
[6] 学術認証フェデレーション：https://www.gakunin.jp
[7] 情報処理推進機構：“「オンライン本人認証方式の実態調査」報告書について”，https://www.ipa.go.jp/security/fy26/reports/ninsho (accessed 2017.09)
[8] 松平 拓也：“大学統合認証基盤における多要素認証について”，平成 26 年度第 2 回学術基盤オープンフォーラム@NII, feb, 2015
[9] tiqr: https://tiqr.org (accessed 2017.09)
[10] YubiKey: https://www.yubico.com/products/yubikey-hardware (accessed 2017.09)
[11] 河野 圭太，稗田 隆，中村 素典：“Shibboleth と OpenAM の連携による認証レベルを制御可能なシングルサインオン基盤の構築”，学術情報処理研究，No.21, pp.71-81, 2017.9
[12] FORGEROCK Access Management: https://www.forgerock.com/platform/access-management (accessed 2017.09)
[13] 日本ネットワークインフォメーションセンター：https://www.nic.ad.jp/ja (accessed 2017.09)
[14] Spring: https://spring.io (accessed 2017.09)
[15] Spring Web Flow: http://projects.spring.io/spring-webflow (accessed 2017.09)