

# 九州工業大学情報基盤センター年報

## 第4号

### 2024.3

#### 目 次

#### 巻 頭 言

九州工業大学における今後の IT 基盤整備と DX について ..... 中村 豊 ... 1

#### 解 説

部局内メールアドレスの Microsoft365 への集約について  
..... 林 豊洋, 黒崎 覚 ... 3

ビデオ会議サービスの提供と API 連携機能の導入  
..... 大西 淑雅, 山口 真之介 ... 11

九州工業大学における脆弱性の検査と改善の取り組み  
..... 佐藤 彰洋, 福田 豊, 中村 豊 ... 25

#### 解 説 (研究紹介)

DGA マルウェアにより生成された悪性ドメインの検出  
..... 佐藤 彰洋, 福田 豊, 中村 豊 ... 31

辞書に基づく DGA マルウェアにより生成された悪性ドメインの検出  
..... 佐藤 彰洋, 福田 豊, 中村 豊 ... 45

#### 報 告

利用実績 ..... 59

教育研究支援 ..... 71

広報出版・セミナー開催 ..... 73

本年度の活動 ..... 75

センター人事異動および職員配置 ..... 78

情報基盤センター規則等 ..... 79

## センターの各種メーリングリスト

名称	用途
support@isc.kyutech.ac.jp	情報基盤センターに関する一般的な質問用
tebiki@isc.kyutech.ac.jp	「オンラインガイド」（教育システム環境用 WWW サーバ上で公開中）に関する質問用

## センターへの連絡

連絡先名称	場所	電話	メール
飯塚利用者窓口	センター棟 (2F)	0948-29-7558	support@isc.kyutech.ac.jp
飯塚事務室	センター棟 (1F)	0948-29-7555	jimu@isc.kyutech.ac.jp
戸畑利用者窓口	情報学習プラザ (2F)	093-884-3471	support@isc.kyutech.ac.jp
戸畑事務室	総合教育棟 (2F)	093-884-3470	jimu@isc.kyutech.ac.jp

◇◇◇◇◇  
巻頭言  
◇◇◇◇◇

## 九州工業大学における今後の IT 基盤整備と DX について

中村 豊<sup>1</sup>

現代は VUCA(Volatility, Uncertainty, Complexity, Ambiguity) 時代と呼ばれる不安定で不透明な先の見通せない時代と認識されています。このような変化に富んだ時代であるからこそ、目前の課題の解決だけに留まらず、様々な価値観を融合させた新たな価値の創造に挑戦することが求められています。情報統括本部では、本学が今後取り組んでいくべき挑戦について、Kyutech-DX ビジョン 2023[1]として取りまとめました。このビジョンを実現するために情報基盤センターでは、ICT 利活用教育研究基盤運用室、ネットワークセキュリティ基盤運用室および DX 推進室とともに以下の様な課題について取り組んでいます。

ICT 利活用教育研究基盤運用室では全学統合 ID システムを更新し、SAML/SSO ならびにアカウント管理を柔軟に対応できるシステムの導入を進めています。これまで社会人に対して公開している学習サービスにおいて、個別の部局でゲストアカウントを発行していましたが、今後は一括管理が可能となり部局の負担を軽減することができます。また、新システムでは Microsoft 365 とのアカウント連携の強化機能を有しているため、これまで以上に Microsoft 365 の利便性が高まることが期待できます。さらに ICT 教育研究基盤システムの仕様策定が開始されており、2PB を超えるキャンパス基幹ストレージの構想や学内プライベートクラウド基盤としての機能が期待されます。DX 推進室との連携では全学的な Teams 電話の導入が始まり、固定電話および内線電話に縛られない業務の実現が期待されます。今後は電子決裁システムの導入を進めることで、業務のオンライン化・高効率化の実現が期待されます。

ネットワークセキュリティ基盤運用室では全学ネットワーク基盤を更新し、より高速で安心、安全かつ安定した運用が可能なネットワーク基盤の構築を進めています。最新規格に対応した学内無線 LAN 網のさらなる拡充や、EVPN-VXLAN を用いた拠点間の L2 延伸機能を積極的に用いることでネットワーク運営の効率化を図ります。さらに 100 Gbps を超える超高速ネットワークを導入し、学内の様々なニーズに対応できる情報通信基盤の整備を進めます。また、学内権威 DNS サーバやメールサーバ、Web サーバの集約、さらにプライベートネットワーク基盤のサービス提供を開始し、学内サービスの集約化を進めています。

今後、さらに IT 基盤の重要性は高まってくると思われれます。システムだけではなく制度・組織等を常に更新し、課題を解決していくことが求められます。これらの課題解決には情報基盤センターだけではなく、教育組織や事務組織との連携が必要不可欠です。情報基盤センターとしては、さらなる DX の推進のために全学一丸となった取り組みを進めていきたいと考えています。今後とも、ご指導・ご協力をよろしくお願いいたします。

---

<sup>1</sup>情報基盤センター長 yutaka-n@isc.kyutech.ac.jp

巻頭言

## 参考文献

- [1] 九州工業大学 Kyutech-DX ビジョン 2023 : <https://www.kyutech.ac.jp/information/kyutech-dxvision.html>

◇◇◇◇◇  
解 説  
◇◇◇◇◇

## 部局内メールアカウントの Microsoft365 への集約について

林 豊洋<sup>1</sup>  
黒崎 寛<sup>2</sup>

### 1 はじめに

様々なコミュニケーションツールが存在する中、電子メールは古くから用いられていることもあり、依然として必要不可欠な手段となっています。インターネット上で一般的に用いられる SMTP を用いた電子メールは、標準化された 1980 年代の設計思想が残っており、小規模の組織（大学においては、研究室や学科等の部局ドメイン単位）においてもメールシステムを構築し、サービスとして公開することが可能であり、それらの組織間でメールの授受が行われています。

対して近年は、メールサービスの維持が困難となる要因が増え、かつ深刻化しています。特に、小規模の組織におけるメールサービスの維持は困難となりつつあり、本学においても例外ではありません。

この状況に対応するため情報統括本部では、部局ドメイン単位で稼働する部局内メールサービスについて、全学サービスである Microsoft 365 テナント（以下、M365）への集約を検討しました。様々な方法から、既存の九工大メールアドレスと部局内で用いられていたメールアドレスを対応付ける方法を確認しました。この方法により、テナントの M365 ライセンスを追加消費することなく、部局で用いられていた個人用メールアドレス、メーリングリスト、事務・講義等で用いられていた共有メールアドレスを M365 上で利用することが可能となりました。

### 2 電子メールサービスを取り巻く状況

利用者視点での電子メールサービスとは、「メールクライアントを使って、自由に電子メールの送受信ができる」ものです。非常に単純な要求であるため、以前はメールサーバ（MTA, MRA）に加え、対象となるドメインの MX レコードを指定する程度でサービスは稼働できました。

しかし近年の電子メールサービスは、多数の情報システムの組み合わせが必要となっています。その要因の一部を後述します。

#### 2.1 求められる機能の多様化

電子メールサービスに求められる機能は、継続的に増加・高度化しています。利用者視点での要求としては、以下のようなものが挙げられます。

**オンライン上にメールを蓄積したままにしたい** 以前は PC 側にメールをダウンロードして読めれば良い様な利用形態でしたが、複数の環境で同様のメールボックスを維持するには、IMAP 等に対応した MRA に加え、大容量のストレージを備える必要があります。

<sup>1</sup>情報統括本部情報基盤センター 准教授 toyohiro@isc.kyutech.ac.jp

<sup>2</sup>情報統括本部情報基盤課情報基盤運用係 係長 kurosaki-s@jimu.kyutech.ac.jp

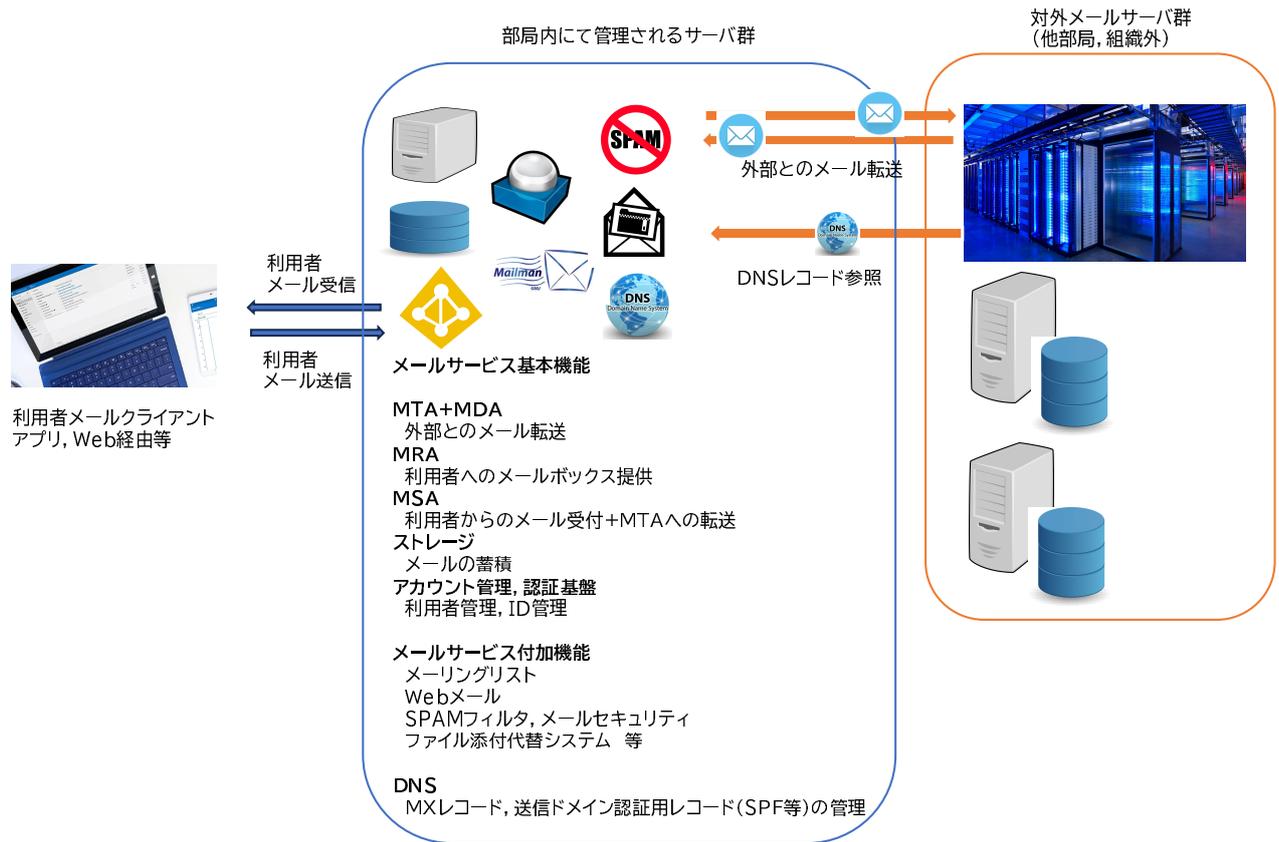


図 1: 電子メールサービスを構成する機能群

**Web メールを使いたい** メールクライアントとして Web ブラウザを用いることにより、場所や環境を問わずにメールの送受信が可能となりますが、Web メールシステムを構築する必要があります。

**メーリングリストを運営したい** メンバー管理機能や、専用の SMTP ヘッダを付与する機能を備えたメーリングリスト管理機能が必要となります。

**SPAM を除去してほしい** 受信したメールが SPAM に該当するか判定し、所定の動作（迷惑メールと判定しヘッダを付ける、フォルダ分けする等）を行う機能が必要となります。

**フィッシング、マルウェア対策をしてほしい** 添付ファイルやメールに記載の URL が安全であるかを判定し、危険な場合は利用者を保護する機能が必要となります。

また、サービスの管理者視点での要求も増加、高度化しています。

**メールクライアントの設定が簡単に出来るようにしてほしい** メールクライアント利用時の設定となる、ログイン ID 情報の入力・サーバへの接続情報の指定は煩雑で、利用者からの問い合わせが多い内容となります。簡単化するには対応した認証基盤が必要となります。

**ファイル添付の代替を備えたい** ファイル添付は機微情報の流出等の要因になるため、オンラインストレージ等の代替システムを備える必要があります。

**メールセキュリティを導入したい** 大量のマルウェア送受信等の挙動を検知し、送受信をブロックする機構等により、セキュリティを高める必要があります。

**ログの保存、検索を効率化したい** メール送受信ログは、不正通信の追跡に必須なものとなります。保存機能に加え、効率的な検索が可能なシステムが必要となります。

このような様々な要求に対応したメールシステムは、図 1 に示すような機能群で構成されることとなり、以前のシステムの構成とは全く異なるものが要求されます。

## 2.2 セキュリティ確保の困難化

前述のように、近年求められるメールシステムを構成するには、多数の情報システムを稼働させる必要があります。情報システムには脆弱性が必ず含まれており、継続的なバージョンアップ・パッチ適用等による脆弱性対策が必要となります。電子メールの送受信用の SMTP サーバでさえ決して枯れた技術になることはなく、定期的に脆弱性が発見されています。

また、以前は問題ないとされていた設定値やシステム構成でさえ、攻撃されるパターンや安全でないと見做されることがあります。各種サーバの設定値の確認や、新たな情報システムや設定を追加し、サービスを継続する必要があります。

多数の情報システムが稼働し、それらの組み合わせによって構成されるシステムであるため、このような継続的な対策は困難を極めます。

## 2.3 小規模の組織における懸念点

特に小規模の組織においては、サービスの管理者が少ない（特に、情報システムの管理を主業務とする職員がいない）事が多いと言えます。セキュリティ確保に加え、入職・理退職に応じたアカウント管理、利用者のトラブル対応などを含めると、メールサービスの維持管理は更に困難となります。

また、メールシステムを構成する情報システムは利用者数に関わらず整備する必要があります。サービスを維持するための設備投資を継続する必要がありますが、小規模の組織ほど、いわゆるコストパフォーマンスは不利なものとなります。

## 3 全学サービスへの集約

情報統括本部においては、重要度の高い情報システムを全学サービスとして管理運営しており、これまで部局内で構築していたシステムの集約化を推進しています。既に、ファイアウォール、プライベートネットワーク、権威 DNS ホスティング等のネットワークシステムについては、集約を進めています [1]。

前節で述べた電子メールサービスを取り巻く状況は本学においても例外ではなく、部局単位でのサービスの維持は困難になりつつあると認識しています。このような経緯より、新たに部局ドメイン単位で稼働する部局内メールサービスについて集約を検討しました。

集約に用いる方法や情報システムは様々考えられますが、SaaS を活用する方法が、多数のメールアカウントを収容でき、利用者・管理者とも豊富な機能を活用でき、かつコストメリットも高いものとなります。特に本学においては、全学のメールサービスである九工大メールサービスに M365 を活用しており、本学での運用に特化した設定がなされたテナント（以下、全学テナント）が稼働しています。

M365 はメールサービスに関わる多数の機能を備えており、部局内で稼働している大半の機能はより利便性が高いものに代替されます。SaaS であるため、脆弱性対策や新機能への追従はシステム側で実施されます。また、全学テナントの運用は ICT 室が担当するため、部局内での運用は不要となります（図 2）。

### 3.1 部局内メールと全学テナント上のサービスとの対応付け

本節では、部局内メールの全学テナント上での展開に関する詳細について解説します。ここでは、部局のメールアドレスを x.kyutech.ac.jp、全学テナント上は九工大メールアドレス mail.kyutech.jp とします。

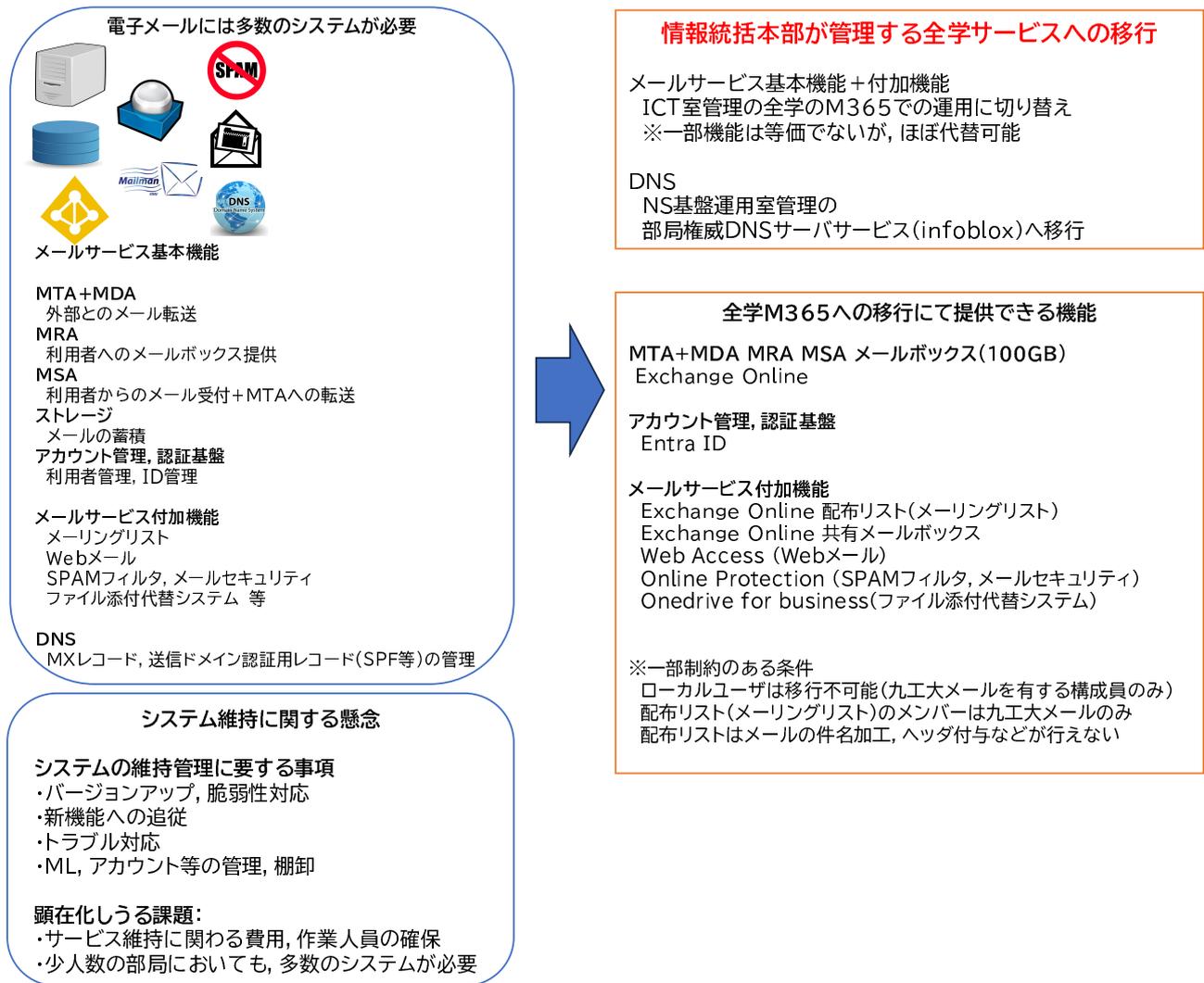


図 2: 部局内メールサービスの全学サービスへの移行

M365は、複数のメールドメインを登録し、それぞれのドメインにおいて任意のメールアドレスを発行することが可能なマルチドメイン機能を有しています。したがって、部局ドメイン x.kyutech.ac.jp を全学テナントに登録することにより、九工大メールと同様の A5 ライセンスを持つメールアドレスを新規に発行することや、メーリングリストに対応する配布リストを作成する事は可能です。

しかし、メールアドレスの新規発行を行うと、M365 ライセンスを消費することとなります。特に、豊富な機能やセキュリティ対策がなされた A5 アカウントは予算規模の関係上、教職員当たり一つの割り当てが現実的であり、部局ドメイン用に追加消費する方法は考慮できません。また、学科等においては「事務室用のメールアドレス」「講義用のメールアドレス」と称されるものがあり、担当者が変わる際にメールアドレスを引き継ぐことや、複数人でメールアドレスを共用する利用形態が存在します。M365 のライセンスは単一の人物に付与することが原則であり（ユーザライセンス）、人物が存在しないアカウントの発行や再利用と判断される利用は行えません。これらの条件を含めて検討を行い、以下の方式で部局内メールと全学テナント上のサービスとの対応付けを行う事を決定しました（図3）。

### 部局内メールアドレスの移行

- 部局内メールアドレスの利用者が、九工大メールアドレス（M365 アカウント）を持つことが条件（持たな

い場合は移行不可)

- 部局内メールアドレスと同一名称の「M365 配布リスト」を作成（部局内メールアドレス名でメール受信可能となる）
- 作成した M365 配布リストのメンバーとして、利用者の M365 アカウントを追加（部局内メールアドレス宛のメールは九工大メールアドレスのメールボックスに着信）
- 利用者の M365 アカウントに対して、部局内メールアドレスと同一名称の配布リスト名を用いて、メール送信可能とする許可を与える（部局内メールアドレス名でメール送信可能となる）

### 部局内メーリングリストの移行

- メーリングリストのメンバー全てが九工大メールアドレスを持つことが条件（外部ユーザを含む場合は移行不可）
- 部局内メーリングリストと同一名称の「M365 配布リスト」を作成（部局内メーリングリスト名でメール受信可能となる）
- 受信対象となる M365 アカウントをメンバーとして追加（部局内メーリングリスト宛のメールは九工大メールアドレスに着信）

### 引き継ぎ，共用メールアカウントの移行

- 事前に配布リストへの移行で運用可能であるか部局に判断いただく
- 共用となる場合は，部局内の対象メールアドレスと同一名称の「共有メールボックス」を作成（対象メールアドレス名でメール送受信可能となる．また，共有メールボックスは M365 ライセンスを消費しない）
- 共有メールボックスにアクセス可能なメンバーとして，利用者の M365 アカウントを追加

これらの対応付けにより，既存の部局内メールアカウント（共用を含む），メーリングリストについて，全学テナント上の M365 ライセンスを消費することなく移行可能となります．利用者は，九工大メールアドレスと移行された部局内メールアドレス双方が利用可能です．

ただし，部局内メールアドレスの利用者が九工大メールアドレスを持たないローカルユーザである場合は移行不可であることや，メーリングリストの移行については，メンバー全てが九工大メールアドレスを持つこと，移行後の配布リストはメーリングリストの機能としては不足がある（件名へ番号を付与する加工機能等がない）ことなど，いくつかの制約を受容いただく必要があります．また，いずれも新規追加，メンバー変更，運用変更等が生じる場合は，都度 ICT 室へ通知頂くことにより対応が必要となります．

## 4 部局内メールサービスの全学テナントへの移行の流れ

部局内メールサービスの全学テナントへの移行に関する時系列的な手順を示します（図 4）．

まずは，移行を想定される日時の数か月（半年程度を想定）前に，部局長等から情報統括本部へ打診頂くことで作業着手となります．作業着手後，実務を担当される部局内情報システムの管理者等から，情報統括本部へ現状のシステム構成・メールアカウントの発行状況（個人用メールアドレス，メーリングリストの構成，共用メールアカウントの有無，九工大メールアドレスとの対応関係）についてご提供ください．その後，統括本部内で移行可能範囲の判断を行い，申請いただいた部局へ移行プランについてご説明いたします．

部局において移行決定の判断を頂いた場合，具体的な移行日について調整を実施いたします．移行日一か月前までに，移行対象のメールアカウント情報に基づき，統括本部内で全学テナントへ側の事前設定を実施いたします．

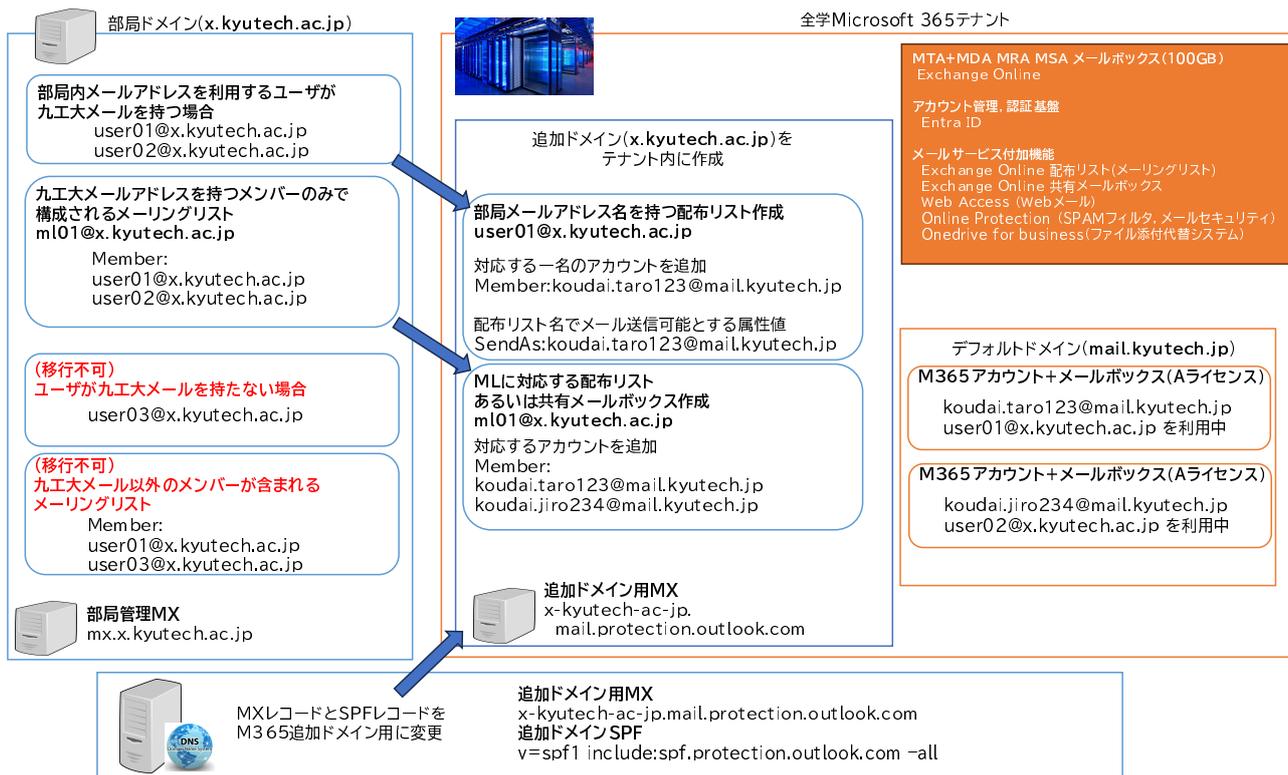


図 3: 部局内メールと全学テナント上のサービスとの対応付け

移行日一か月前に、部局メールアドレス (x.kyutech.ac.jp) を全学テナントへ追加いたします。この際、部局メールアドレスを用いた M365 テナントが既に存在する場合、部局内情報システムの管理者と協調し、ドメインの移行作業が必要となります。

ドメイン追加完了後、移行日に設定頂く DNS レコードの情報を通知、移行日当日に全学テナントへ移行対象の部局内メールアドレスの追加を実施いたします<sup>3</sup>。

送受信テストを以て、全学テナントへの移行作業は完了となりますが、部局内に存在する既存 MTA をはじめとするシステム群の運用停止、既存メールボックスからのメール移行作業については、部局にて対応頂くこととなります。

## 5 まとめ

本稿では、情報統括本部にて提供可能となった部局内メールアドレスの全学テナントへの移行・集約について解説しました。

メールサービスは要求される機能が増え、多数の情報システムで構成する必要があるため、管理コストの増大や継続的なセキュリティ対策を要することが課題となりつつあります。この状況に対応するため情報統括本部では、部局内メールアドレスを全学テナントへ追加し、部局内メールアドレスと九工大メールアドレスを対応付ける方法により、メールサービスの移行・集約を可能としました。

2023 年度に 1 部局の移行が実施され、更に 2 部局について 2024 年度の移行を検討しています。メー

<sup>3</sup>事前に移行対象のメールアドレスを追加した場合、配布リストはトランスポートルール適用対象外となり [3]、全学テナントから移行前の部局内メールサービスへメールが送付されない (追加した配布リストに到達してしまう) ため、当日部局内情報システムの管理者と協調し実施することとなります。

ルサービスの維持が困難となりつつある部局におかれましては、本サービスの利用を検討頂けると幸いです。

## 謝辞

本件の初事例となりました大学院工学研究院物質工学研究系内情報システムを統括される齋藤泰洋准教授には、移行手順に関する試行実験や配布リストを用いたメール送受信に関わる検証、学科内の調整等を進めて頂きました。深く感謝申し上げます。

## 参考文献

- [1] 九州工業大学情報統括本部, "学内各種サービス", <https://www.kiban.kyutech.ac.jp/service.html>, (2024年1月31日参照)。
- [2] 嶋吉隆夫, 笠原義晃, 平川新, 亀岡謙一, 平野広幸, 藤村直美, "九州大学における組織別運用メールサービスのクラウド集約への取り組み", 大学ICT推進協議会2020年度 年次大会, FB2-2, 2020.
- [3] Microsoft Learn, "Exchange Online のメール フロー ルールでの条件と例外", <https://learn.microsoft.com/ja-jp/exchange/security-and-compliance/mail-flow-rules/conditions-and-exceptions>, (2024年1月31日参照)。

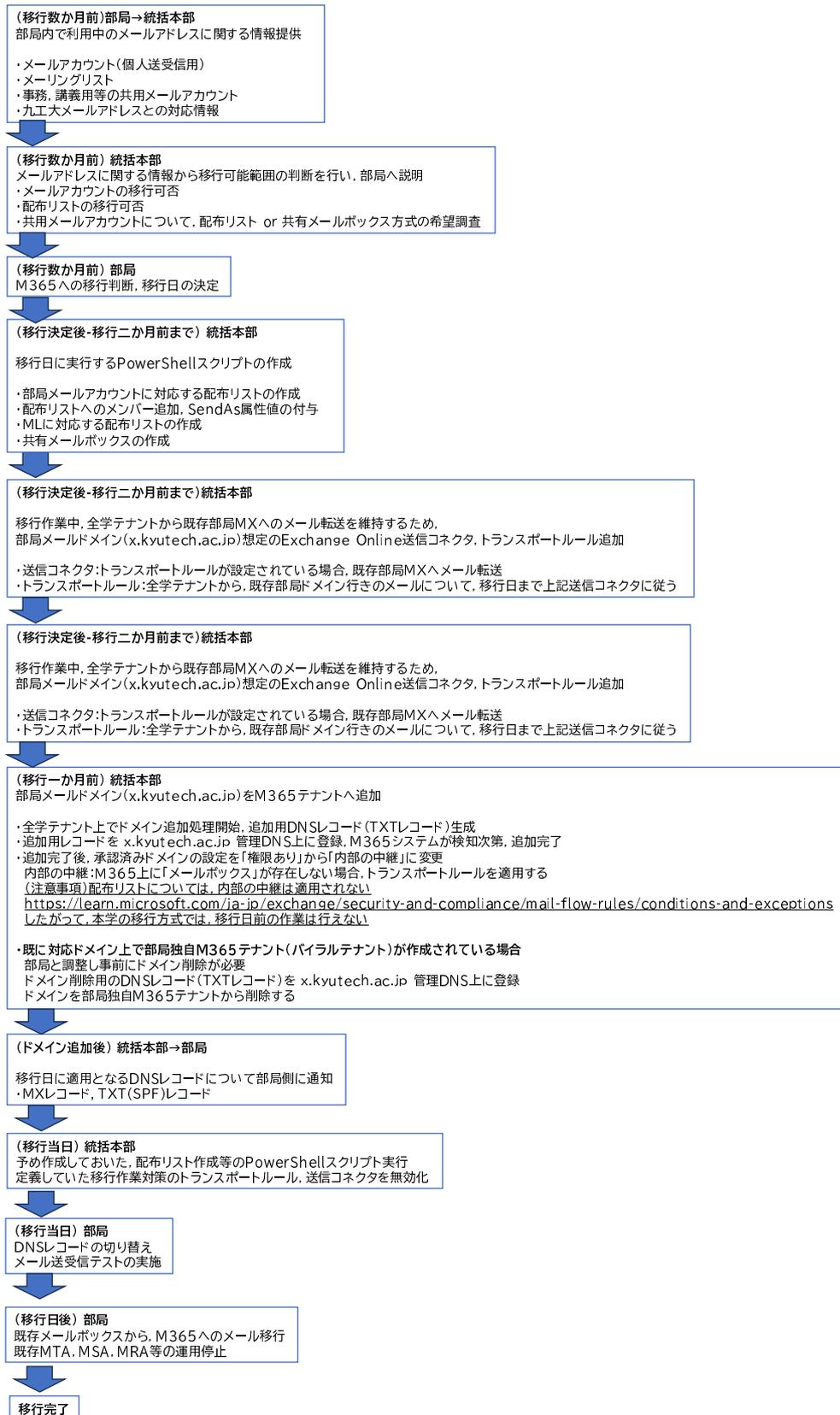


図 4: 部局内メールサービスの全学テナントへの移行の流れ

◇◇◇◇◇  
解 説  
◇◇◇◇◇

## ビデオ会議サービスの提供と API 関係機能の導入

大西 淑雅<sup>1</sup>  
山口 真之介<sup>2</sup>

### 1 はじめに

1986年の情報工学部の設置に伴い、九州工業大学は戸畑キャンパスおよび飯塚キャンパスの2キャンパス体制<sup>3</sup>となった。これに伴い、およそ41km離れたキャンパス間において、ビデオ会議システムを用いた会議が実践されてきた。2005年以降、ビデオ会議システムを用いた教育実践も数多く行われ、サテライトキャンパスを含む複数の教室で同期型の遠隔授業なども実践されてきた [1, 2]。この頃のビデオ会議システムは専用機の形態をとっており、利用者にとって分かり易いインタフェースを持つ反面、設置教室に限られるなどの課題があった。そこで、専用機をラックにまとめ、複数の教室で運用可能な工夫 [3] など行ってきた。

一方、2014年からキャンパスネットワークを管轄する、情報基盤運用室（現：ネットワークセキュリティ基盤運用室）が活動を開始したことで、キャンパス内の無線LAN環境の整備が計画的に行われるようになった。このことは、2016年度頃から情報科学センターを中心に、議論を開始した「ノートパソコンの必携化」への移行を後押しし、キャンパスネットワークをより積極的に活用する方向性が示された。その結果、2018年度の一部学部の一試行を経て、2019年度の入学生から全学的にノートパソコンの必携化 [4, 5] を開始した。また、キャンパスネットワークの安定性の向上と無線LAN環境の充実、専用機スタイルのビデオ会議システムや多地点接続システム (MCU) にも影響を与え、学習教育センターでは Adobe Connect Pro (2013年6月～2019年3月) の試行を踏まえ、より汎用性の高いビデオ会議サービス (WebEX 2018年10月～) の活用を模索してきた。

このような背景の中、2020年初旬に、新型コロナウイルス感染症の流行が拡大し、多くの教育機関においても、遠隔教育を実施することになった。九州工業大学においても、「遠隔授業対応支援WG」が設置され、部局間連携や教職員の協力により遠隔授業を実践してきた。この経験は、Zoom や WebEX に代表されるビデオ会議サービスや学習支援サービス (Moodle) を用いた教育手法を、多くの教職員に浸透／定着させる結果となった。

本稿では、学習教育センターが2020年4月末より提供／管理してきた、ビデオ会議サービス (Zoom) を用いた教育／研究活動について、概ね3年間の活用状況についてまとめる。また、Zoom Education Site ライセンスに含まれる、Zoom API の活用例を紹介すると共に、ビデオ会議サービス (Zoom) のクラウドレコーディングにおける、管理・運営上の工夫について報告する。

### 2 遠隔授業

2020年4月15日に当時の教育担当理事から「遠隔授業対応支援WG設置の提案」がなされ、「学生教育においても、当の間、3密条件（換気の悪い密閉空間、多数が集まる密集場所、間近で会話や発声を

<sup>1</sup>情報基盤センター 准教授, <https://www.isc.kyutech.ac.jp/>, 学習教育センター兼任

<sup>2</sup>学習教育センター 助教, <https://www.ltc.kyutech.ac.jp/>

<sup>3</sup>2024年3月現在は3キャンパス

する密接場所)を満たす状況を作らないことが求められます。そのため、遠隔授業を中心に一部対面授業を併用する形態や、全コマを遠隔授業で実施する形態が、不可欠になってきました。」の趣旨に従って、「実際に遠隔授業を行う教員(と学生)の支援体制の構築」に相当する「遠隔授業対応支援WG」が2020年4月16日に結成(図1)された。

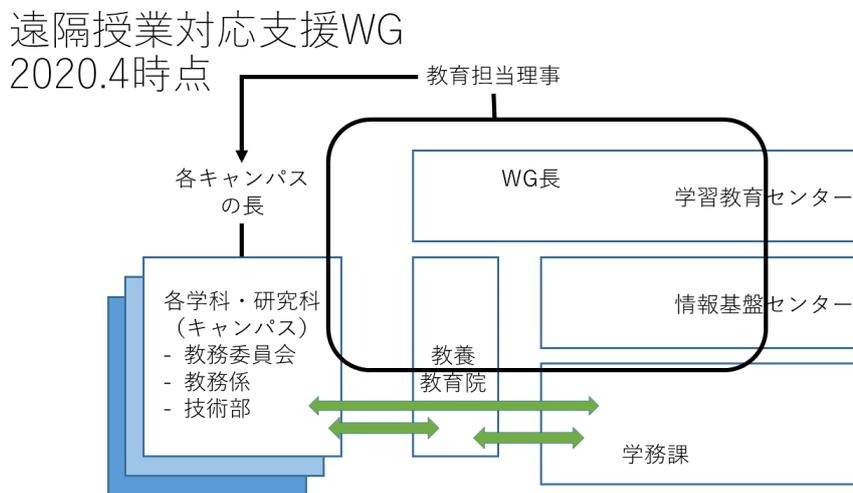


図1: 遠隔授業対応支援WG(令和2年度遠隔授業支援WG(全学)資料より)

2020年3月頃にZoom社が「Zoom無料ベーシックプラン」の制限解除を表明<sup>4</sup>したことにより、遠隔会議や遠隔授業を行うツールとしての先行テストが、ボランティア教職員によって行われた。その評価結果の高さや他大学の動向・関連する情報の多さから、九州工業大学においてもビデオ会議サービス(Zoom)を導入した[6]。学習教育センターでは、認証基盤[7]を新に構築し、多くの教職員・学生が本サービスを使った、教育/研究や学習活動を実践できるように工夫した。また、学生間の交流やグループ学習を支援できるように、学生にもホストライセンスの付与を試みた[8]。

2020年度は、遠隔授業を実践しつつ必要ライセンスの購入と配布、その管理体制の構築を進めた。表1に2020年4月以降の契約ライセンス数の推移を示す。保有ライセンス数が十分でなかったため、必要時期に応じて、Zoomホストのライセンス有とライセンス無を切り替えていたが、クラウドレコーディング上のファイル管理に課題があり、2021年度からは、サイトライセンスに切り替えた。

表1: Zoomライセンスの導入状況(個別契約を除く)

ライセンス数契約	合計	事務用	授業用	学生用	備考
2020.4.30~2021.4.29	396	38	360	-	
2020.7.01~2021.4.29	406	10	0	-	
2020.8.01~2021.4.29	427	1	20	-	
2020.9.01~2021.4.29	437	10	0	-	
2020.10.01~2021.4.29	497	30	30	-	
2020.11.01~2021.4.29	700	31	152	20	授業補助者を含む
サイトライセンス契約	合計	基準数	職員数	学生数	備考
2021.4.30~2022.4.29	6536	908	908	5628	
2022.4.30~2023.4.29	6571	952	952	5619	大規模5, Webinar1
2023.4.30~2024.4.29	6571	955	955	5616	大規模3, Webinar1
2024.4.30~2025.4.29	確定前	*854	1012	5595	大規模2, Webinar1

\*2024年度の職員数はZoom社のカウント方法変更後の計算値。

※サイトライセンスに一部個別契約を追加。

<sup>4</sup>新型コロナウイルスの感染拡大により休校などの影響を受けた幼稚園から高校までの指定校は、2021年7月31日まで無料のZoomベーシックアカウントのミーティングの40分間の制限が解除されます。

### 3 利用状況

図2に2020年3月から2023年2月までのビデオ会議サービス (Zoom) の利用数を示す。利用数は、Zoom 管理者の「過去のミーティング」を用いて取得したデータから算出した。但し、2020年9月から12月までと2022年4月から7月までのデータは、取得できていないため割愛した。2023年度の利用数を2021年度と比較すると、平均67.2%にとどまっている。これは原則として対面授業とした影響と思われる。

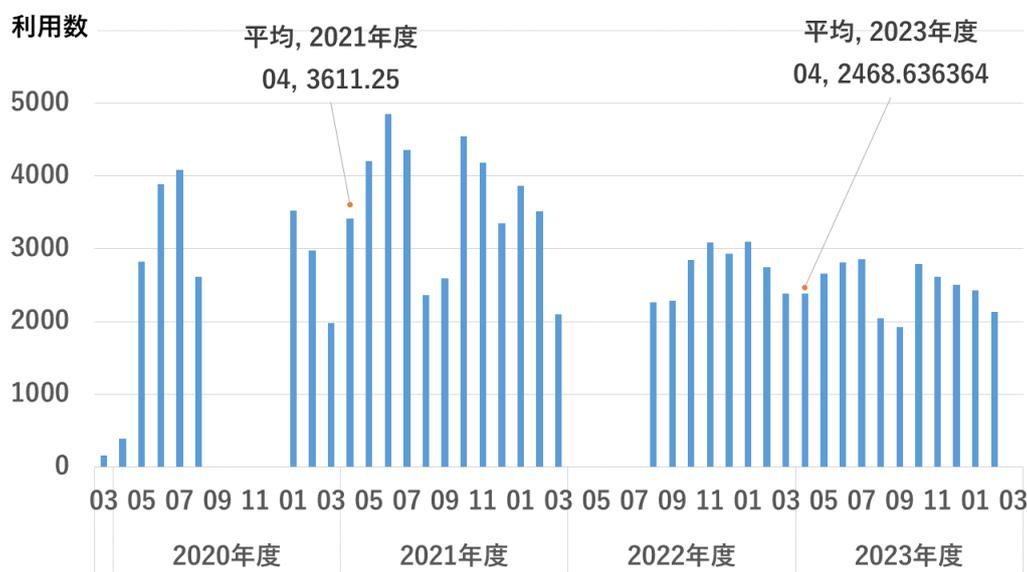


図2: ビデオ会議サービス (Zoom) の利用状況※ 2020.9-12, 2022.4-7 はデータ無

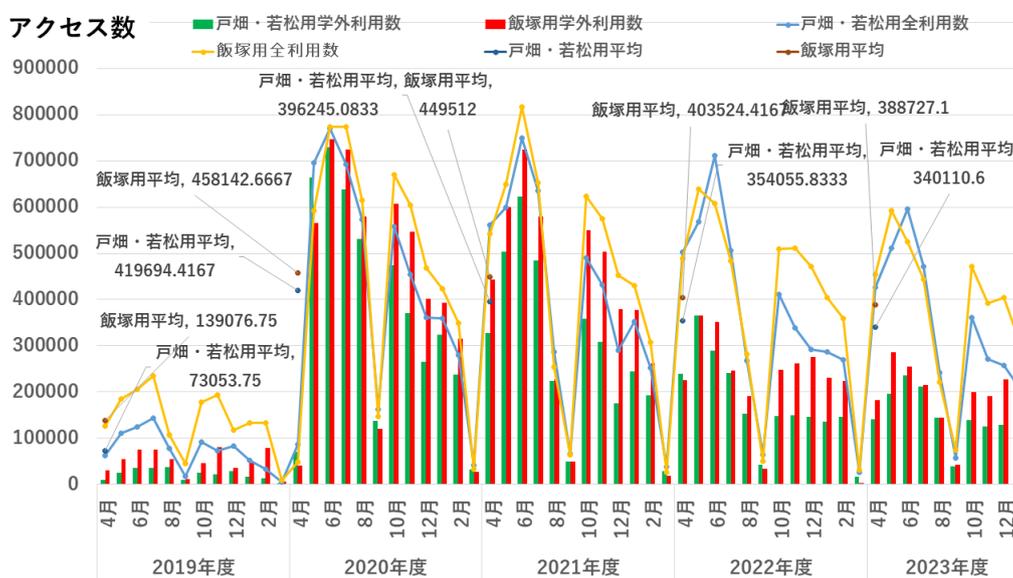


図3: 学習支援サービス (Moodle) の利用状況

一方、学習支援サービス (Moodle) の利用状況 (図3<sup>5</sup>) について、2021年度平均と2023年度平均を比較すると、飯塚用で86.4%、戸畑／若松用で85.8%となっている。つまり、学習支援サービス (Moodle)

<sup>5</sup>アクセス数は、各種リソースの参照回数に加え、各種活動の実施数 (例：課題にファイルを提出した、クイズに回答した、フォーラムに投稿した、投票に回答した、など) を合計したものである。

の利用減少に比べて、ビデオ会議サービス (Zoom) の利用減少が大きくなっていることがわかる。これは、対面授業時の学習支援サービスの利用が、2019年度より定着したことを示している。

### 3.1 ホスト数と活用頻度

現在の Zoom ホスト登録数を表 2 に示す。なお、既に本学所属でない方も含まれている<sup>6</sup>。表 2(左)の「学習教育センター管理」のホストでは、学習教育センター認証基盤 [7] を用いてユーザを管理している。ドメイン el.kyutech.ac.jp 欄のホスト数が、本学の教職員および学生の主たる利用数となる。表 1 の備考に記載の「大規模 (Large Meeting500)」や「Webinar(Webinar500)」は、300 人を超える会議が必要な方に、割り振る運用を行っている。利用目的と利用期間を申請して頂き、ライセンスを追加する形をとっている。なお、3 年間の運用でライセンス数が不足したことはない。

一方、事務系 (情報基盤課) が管理する Zoom ホストは、利用部局が有するメールアドレスを用いた認証であり、主として事務系の会議や各部局のイベントなどで利用する数となる。なお、事務系の一部のホストアカウントについては、オンプレミス運用 [9] (2024 年 3 月末現在で 5 つ) を行っている。

表 2: Zoom ホスト数 (2024 年 4 月 3 日現在※個別契約を除く)

学習教育センター管理		情報基盤課管理	
管理ドメイン	ホスト数	管理ドメイン	ホスト数
el.kyutech.ac.jp	1829	ccr.kyutech.ac.jp	3
gce.kyutech.ac.jp	12	cntl.kyutech.ac.jp	1
gce.kyutech.ac.jp(Rooms)	5	isc.kyutech.ac.jp	1
ltc.kyutech.ac.jp	3	jimu.kyutech.ac.jp	185
その他	3	jimu.kyutech.ac.jp(Rooms)	1
		kiban-i.kyutech.ac.jp	3
		lai.kyutech.ac.jp	3
		life.kyutech.ac.jp	1
		lsse.kyutech.ac.jp	5
		mail.kyutech.jp	3
		tech-i.kyutech.ac.jp	1

\*2024 年 4 月 3 日値の更新済み。

2022 年 7 月から 2024 年 4 月初旬までにおいて、学生のホスト申請数 (ビデオ会議サービスを一度でも利用しようと試みた人) は 475 であった。教職員については、申請制ではなく、本学所属が確認された時点で、ホストが有効化されライセンス (Zoom Education Site ライセンス: 300 人までのミーティングが可) が付与される。なお、教職員のホスト数は 1363 であった<sup>7</sup>。

学習教育センターが有効化したホストアカウントの活用状況を簡易に把握するために、2024 年 3 月末頃のホストアカウントのログイン状況を調査した (図 4)。横軸は、最新の利用が何日以内であるかを示し、縦軸は対象のホスト数を表している。例えば、60 日以上で 90 未満の間でログインがあったホスト数は、職員が 31 で学生が 27 である。なお、折線グラフは、すべてのホスト数 (学習教育センターおよび情報基盤課が管理) の活用分布を示している。但し、360 日以上 of ホストアカウントのログインが記録されていないものは図 4 から除外<sup>8</sup>した。

利用頻度の高い方は図 4 の「0~7:0 以上 7 未満」に記録されている。年度末の計測のため若干の誤差が生じる可能性もあるが、1 年以内のログイン (0~360) があった職員ホスト総数が 686 であるため、概ね 38.2% 方が頻繁に活用していることが読み取れる。同様に、1 年以内 (0~360) のログインがあった学

<sup>6</sup>ホストアカウントの整理はクラウドレコーディングにも影響するため、2024 年 4 月現在において削除や移行は、原則行っていない。認証基盤側での利用制限や無効化などを実施して利用者管理を行っている。

<sup>7</sup>表 2 のホスト数で若干の差異があるが、これは、特例「一人の方が複数のホストアカウントを希望する」やクラウドレコーディング上の保持などの数が影響している。

<sup>8</sup>職員 677, 学生 70, 総数 853 であった。

生ホスト総数が405であるため、概ね13.6%方の活用がわかる。職員は、研究室打ち合せや会議などの活用が高いと思われるが、学生についてはどのような活用が行われているかさらに調査が必要である。

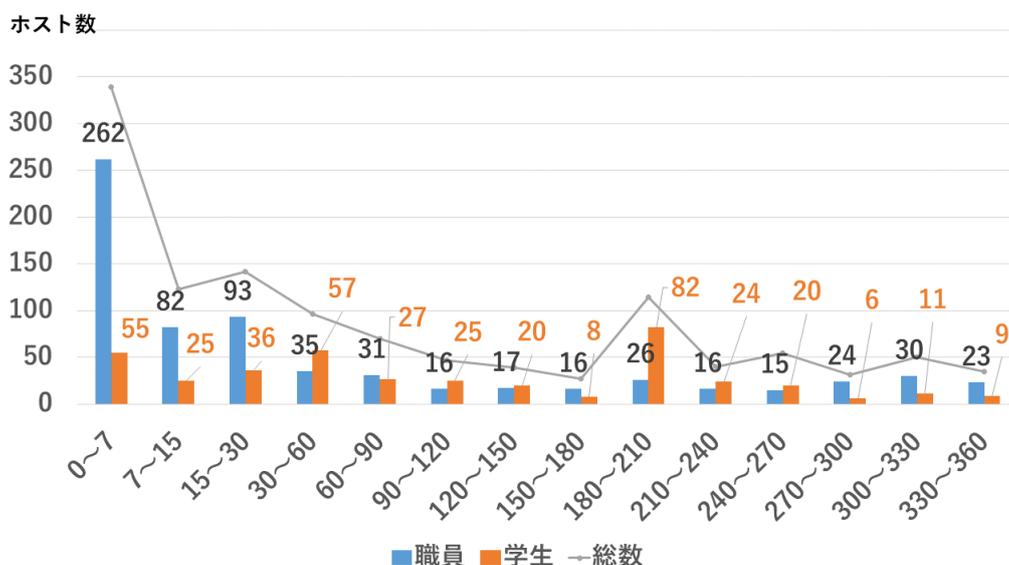


図4: 2024年3月末ごろのホストアカウトの利用状況

### 3.2 クラウドレコーディングの活用

教職員が利用できるクラウドレコーディングの利用状況を図5に示す。なお、本データはクラウドレコーディング上に保持されている保存データ(2024年3月頃のホストユーザ数1723<sup>9)</sup>から算出したものである。

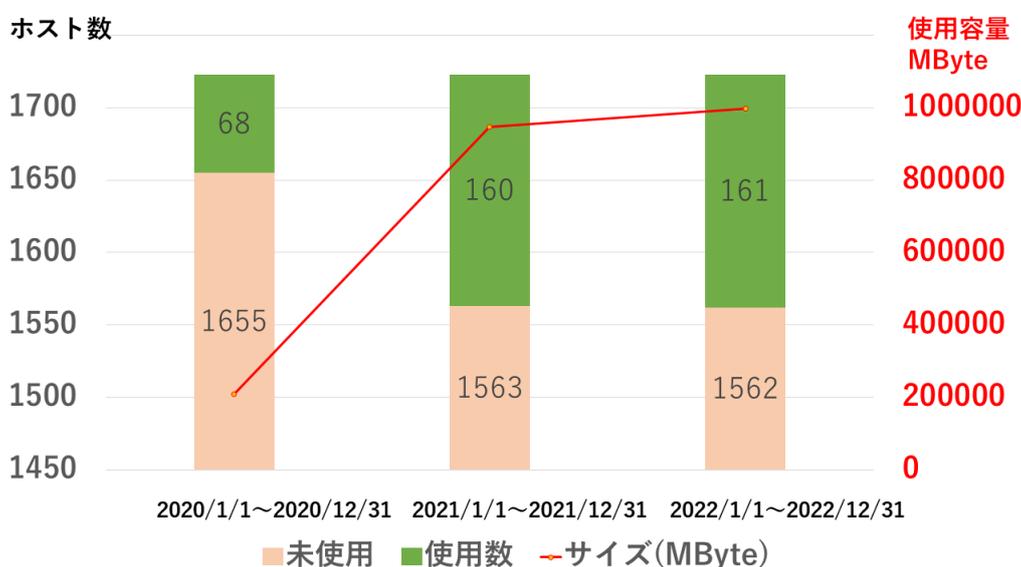


図5: クラウドレコーディング (Zoom) の利用状況

2000年の保存データ数は68ホスト(2024年3月末)となっている。これは、保存可能な容量が少かったため、「授業以外のデータは削除」を依頼<sup>10)</sup>したためと思われる。一方、2021年および2022年の保存

<sup>9)</sup>2024年4月3日現在のホストユーザ数は、教職員1363、学生475、その他

<sup>10)</sup><https://www.ltc.kyutech.ac.jp/10570/>

データ数はほぼ同じで、約 160 ホスト (2024 年 3 月末) 程度であった。こちらは、削除依頼や強制削除を行っていないため、ホストユーザ自身が削除していない(あるいは残した)数となる。2021 年および 2022 年の結果から、クラウドレコーディングを使う教職員の方は、翌年も引続き利用していることが推察できる。また、図 4 のアクティブなホスト数 262(0~7) や 82(7~15) あたりを考慮すると、およそ 4 割の方がクラウドレコーディングを使用していると思われる。

図 6 に 3 年連続 (2000 年から 2022 年まで) クラウドレコーディング上にデータが保存されている方の保存容量を示す。保存容量が多い方の Meeting 情報を抜粋して確認すると、ある教員は、前期に週 2 コマ講義・週 1 コマゼミ、後期に週 2 コマ講義・週 2 コマ実験に対してクラウドレコーディングが行われていた。概ね担当科目数は増減しないことから、授業数はかわらない。しかし、録画ミーティング数は、570(2020 年)、184(2021 年)、130(2022 年) であった。2020 年は対面授業や対面ゼミなどがなかったため、保存数が多くなったと思われる。一方、図 6 に示す方の多くが、保存容量は 2020 年に比べて 2021 年・2022 年が多い傾向が読み取れる。この理由も 2020 年 7 月末に行った「授業以外のデータの削除」を依頼したためと思われる。

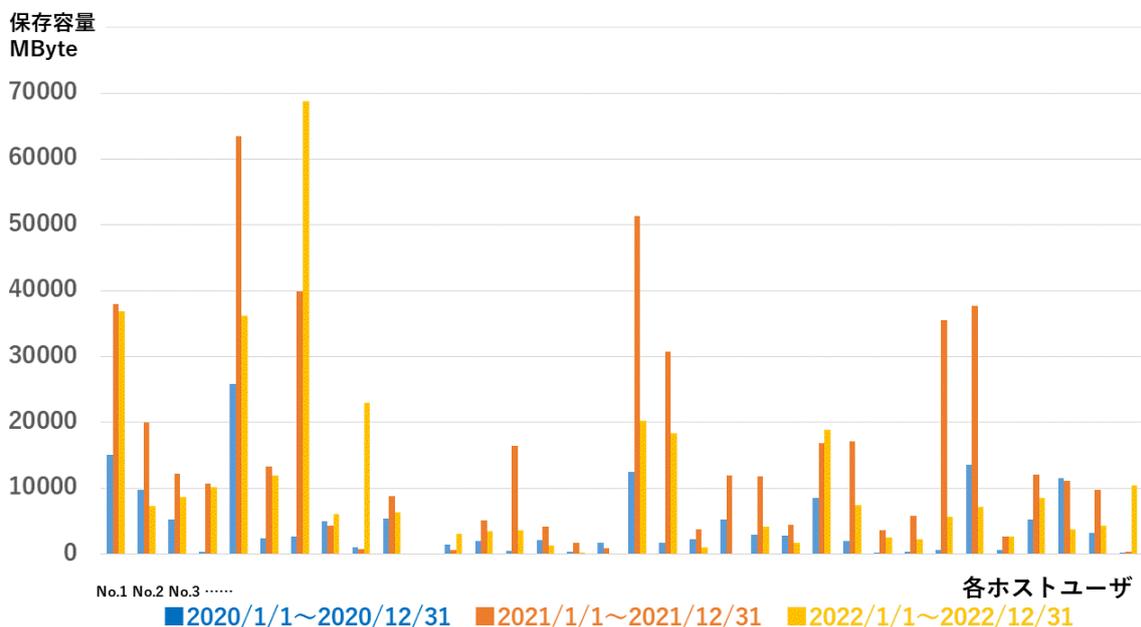


図 6: 3 年連続の利用がある利用者別の利用状況 (MByte)

## 4 Zoom API による保存データの管理

ビデオ会議サービス (Zoom) におけるクラウドレコーディングは利便性が高く、教職員の再利用率も高いため、クラウドレコーディングの容量をどのように確保するかが課題となる。本学の Zoom Education Site ライセンスでは、概ね 3.2TB ほど確保されていたが 2023 年 11 月頃には残容量が少なくなった。そこで Zoom API を用いて録画データを別のファイルサーバに移動させることで残容量の確保を試みた。

### 4.1 設計

ビデオ会議サービス (Zoom) では、様々な情報がインターネット上に公開されている。例えば、Zoom の開発サイトでは様々な質問やその回答が公開されている。クラウドレコーディング上の録画ファイルを取得に関しては「How can I get cloud Recordings video using zoom Api」[10] が Zoom developer サイト

に投稿されている。また、Ricardo Rodrigues さんの GitHub 上に zoom-recording-downloader [11] が公開されている。本学も、この情報を元に録画データの管理を試みた。

図7に Zoom API 連携の流れと保存ファイルの移動を示す。本学の学習支援サービス (Moodle) では、動画ファイルを直接アップロード<sup>11</sup>するのではなく、ストリーミングサービス経由での動画ファイルの配信を推奨している。学習支援サービス (Moodle) では、有償プラグイン HSVIDEO [12] を活用している。HSVVIDEO では、ストリーミングサービスを指定できる。そのため、「動画データアップロードサービス」(既存オンラインストレージ NextCloud) をオンプレミスで構築し、教職員が動画データをストリーミングサービスに簡単に登録できる仕組みを構築 [13] していた。

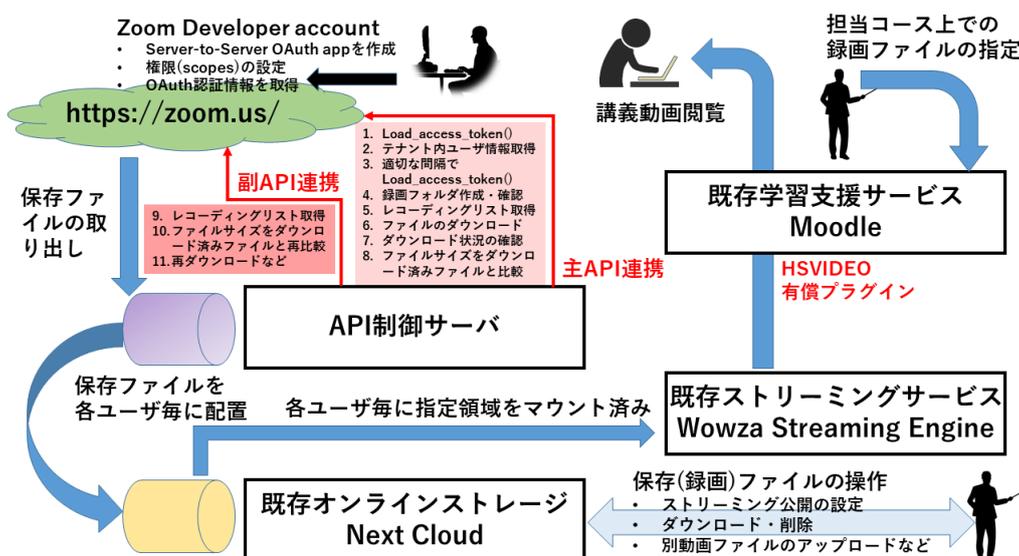


図7: Zoom API による保存ファイルの連携

そこで、図7に示すように、API 制御サーバを新たに設置し、クラウドレコーディングの容量を軽減させることにした。もちろん、教職員が保存データを <https://zoom.us/> から手動でダウンロードし、適切な編集を行って、動画データアップロードサービスに登録することもできる。なお、ビデオ会議サービス (Zoom) のレコーディングに関する、本学の主なグローバル設定(ロック無)を表3に示す。

表3: クラウドレコーディングの主な設定値

項目	値
クラウドレコーディングの共有を許可する	ON
自動レコーディング	OFF
クラウドレコーディング ダウンロード	ON
共有されているクラウドレコーディングにアクセスするにはパスワードが求められます	ON
ホストはクラウドレコーディングを削除できます。	ON
指定された日数が経過した後、クラウドレコーディングを削除します	OFF
ゴミ箱から削除されたクラウドレコーディングの復元を許可します	OFF

## 4.2 ダウンロード機能

Ricardo Rodrigues さんの github 上に zoom-recording-downloader [11](2023年8月頃取得) をベースに本学独自の改良を行った。オリジナルの Python スクリプトでは、load\_access\_token() 関数を最初に行い、

<sup>11</sup>2024年3月現在、サイトアップロード上限値は2GBを設定し、学生等の最大アップロードサイズは500MBである。これは、コースのバックアップおよびリストア時に不具合がないための配慮である。多くの教職員(コース上の教師)は10MB以下の設定を行っている。動画レポートの提出させるような使い方はまれのようである。

その後、レコーディングリストの取得、ファイルのダウンロードを行う手順であった。テスト段階では問題なく取得ができたが、本番ダウンロードでは途中で停止することがあった。オリジナルのスク립トは、ダウンロードが終わったものはファイル (completed-downloads.log) に記録され、次回スク립トの起動時に、再ダウンロードを行わない工夫がされていた。しかし、本学のクラウドレコーディングのヘビーユーザに対しては、複数ファイルのダウンロード中にエラーの発生が確認できた。

そこで、改めて Zoom の開発者サイトを確認した。Server-to-Server OAuth[14] によれば、次のような記載があった。つまり、最初の load\_access\_token() 関数の有効時間は 1 時間であるためスク립トの実行時間が 1 時間以上有する場合は、load\_access\_token() 関数を再度行う必要がある。本学の主 API 連携では、簡易な解決方法として、ホストアカウントの処理が 25 回行われる毎に、load\_access\_token() 関数を再実行するように改良した。

---

Generate access token

Use the grant type to generate an access token. The features of this grant type are:account\_credentials

The token is the owner's access token.

The token's time to live is one hour (3600 seconds).

There is no refresh token.

You can generate and use multiple access tokens.

Tokens stop working when the app is deactivated.

Server-to-Server OAuth apps can be deleted.

Account administrators authorize the scopes available to developers building these app types.

---

また、オリジナルのスク립トでは、会議名に開始日時を付加したディレクトリが作成され、その中に保存されたファイル群がダウンロードされる。ファイル名は def format\_filename(params): で定義されているが、以下のような形式が採用されていた。

ファイル名            保存開始時間 - 会議名 - 保存タイプ - レコーディング ID. 拡張子  
ディレクトリ名        会議名 - 保存開始時間

しかし、日本語の会議名を含むディレクトリやファイル名はストリーミングサービス時のパス指定に不安が残る。オリジナルのスク립トでは、ディレクトリ名およびファイル名として、

教育××学 (2022) 複数コース (13) 後期曜日時限 01 - 2023.01.01 - 07.11 AM UTC  
2023.01.01 - 07.11 AM UTC - 教育××学 (2022) 複数コース (13) 後期曜日時限 01 - Shared Screen With Speaker View - 3127AAAA-60ed-41bc-bbbb-xxxxxxxxxxxxx.mp4

のような命名となってしまう。そこで、会議名からスペースや全角空白を削除し、半角 80 文字を越える場合は、会議名を短くする処理を加えた。あわせて、次のような形式に変更した。

ファイル名            保存開始時間 レコーディング ID. 拡張子  
ディレクトリ名        ミーティング ID 保存開始時間  
シンボリックリンク名 保存開始時間 会議名. 保存タイプ. 拡張子

これにより、ディレクトリ名およびファイル名は次のようになり、さらに追加で、シンボリックリンク名として「教育××学(2022)複数コース(13)後期曜日時限01.SharedScreenWithSpeakerView.mp4」が追加される。動画データアップロードサービス上では、シンボリックリンク名を見ることで保存内容を推察できるようにした。

3Aw+KidCxxxxxxxxcEQVGdWg==2023.01.01.07.11UTC  
2023.01.01.07.11UTC3127AAAA-60ed-41bc-bbbb-xxxxxxxxxxxxx.mp4

なお、オリジナルのスク립トでは、total\_size = int(response.headers.get("content-length", 0)) により

ダウンロードの状況を表示するだけであったが、ダウンロード後のファイルサイズとの比較も念のため行うように改良した。但し、vtt や txt ファイルのような小さなサイズのファイルの場合は、total\_size が取得できなかった。よって、サイズ検証は動画データのようなサイズの大きいファイルのみとした。

その他の改善としては、保存リストの取得(下記参照)を取得済みホスト名でループ処理しているが、本学では、保存リストの重複が確認された。具体的には、同じレコーディング ID であるが異なるダウンロード URL がいくつか確認できた。どちらのダウンロード URL を使っても同じ保存データが取得できるため、レコーディング ID の重複があった場合は、別途ログを残しつつ、重複分のダウンロードをスキップする工夫を加えた。

---

```

for start, end in per_delta(
    datetime.date(RECORDING_START_YEAR, RECORDING_START_MONTH,
        RECORDING_START_DAY),RECORDING_END_DATE,datetime.timedelta(days=30)
):
    post_data = get_recordings(email, 300, start, end)
    response = requests.get(
        url=f"https://api.zoom.us/v2/users/email/recordings",
        headers=AUTHORIZATION_HEADER,
        params=post_data
    )
    .....

return recordings

```

---

### 4.3 チェックモード

ダウンロード機能の開発により、クラウドレコーディング上から1年単位のファイルダウンロードを行っても、エラーが発生することは少なくなった。しかし、ダウンロードに不具合がまったくないことはまれであるため、ダウンロードは1週間程度の期間を設けて、ダウンロード機能のスク립トを複数回実行し、エラーがないかチェックを行った。

一方、ダウンロードされたファイル群の検証も、念のため行えるようにダウンロード機能のスク립トにチェックモードを追加した。具体的には、図7に示すように、「レコーディングリスト取得」後、ダウンロード行わずファイルサイズ「total\_size = int(response.headers.get("content-length", 0))」のみを取得し、「ダウンロード済みファイルとの再比較」を行えるようにした。ダウンロード機能でも同様の比較を行っているが、最終確認のために再比較を行うこととした。ファイルサイズ(Zoom上のサイズとダウンロード先のサイズ)が一致しない場合は、対象データの再ダウンロードを行えるようにし、すべてのダウンロード済みファイルのサイズが一致していることを確認できるようにした。

サイズが一致しているログの例

```

OKSize:check(18/1861)(37:1824),File:2022.08.25.04.01UTCxxxxxx.mp4,
ZoomSize:6612175641,DownSize:6612175641,
Path:/el.kyutech.ac.jp/2022/yyyyyy.edu/bbbbbb==2022.08.25.04.01UTC/2022.08.25.04.01UTCxxxxxx.mp4

```

サイズが一致していないログの例(再ダウンロードが必要)

```

Error :check(19/1862)(37:1825),File:2022.08.25.04.01UTCxxxxxm4a,
ZoomSize:258424752,DownSize:59,
Path:/el.kyutech.ac.jp/2022/yyyyyy.edu/bbbbbb==2022.08.25.04.01UTC/2022.08.25.04.01UTCxxxxxm4a

```

## 4.4 API連携の注意点

以上、Ricardo Rodrigues さんの zoom-recording-downloader をベースに本学独自の改良について説明した。Zoom developers サイトには、API 関係の詳しい情報が公開されているが、実際の動作する zoom-recording-downloader は非常に参考になった。特に、スコープ追加に不具合があると検証に時間がかかる原因となるため、よくチェックすることをお勧めする。参考までに、実際に本学で設定したスコープを表 4 に示す。

なお、load\_access\_token() 関数の自動呼び出しについては、GitHub の精神に従ってコメントを追加したいと考えている。

表 4: 実際に本学で設定したスコープ一覧

Scope Name	ID
View and manage sub accounts	\account:master
View account info	\account:read:admin
View and manage account info	\account:write:admin
View information barriers	\information_barriers:read:admin
View information barriers	\information_barriers:read:master
View and manage information barriers	\information_barriers:write:admin
View and manage information barriers	\information_barriers:write:master
View and manage sub account's user meetings	\meeting:master
View all user meetings	\meeting:read:admin
Get a meeting's encoded SIP URI	\meeting:read:admin:sip_dialing
View and manage all user meetings	\meeting:write:admin
Get CRC dial string with passcode	\meeting:write:admin:sip_dialing
View live streaming meeting token information	\meeting_token:read:admin:live_streaming
View local archiving meeting token information	\meeting_token:read:admin:local_archiving
This scope allows an app to view an account's users' local recording meeting token information	\meeting_token:read:admin:local_recording
View and manage sub account's user recordings	\recording:master
View all user recordings	\recording:read:admin
View and manage all user recordings	\recording:write:admin
View and manage sub account's user information	\user:master
View all user information	\user:read:admin
View users information and manage users	\user:write:admin

<https://github.com/ricardorodrigues-ca/zoom-recording-downloader> の usage を参照して設定。

また、Zoom の契約によっては、API 呼び出しの制限値が異なる。よって、Zoom developers サイトの Rate limits[15] も一度チェックすることをお勧めする。

## 5 保存データの利用者への引き渡し

クラウドレコーディングから引き上げた保存ファイル群は、ホストアカウントの保有者にアクセスできるようにする必要がある。先に説明したとおり、本学では、ストリーミングサーバへの動画登録を可能とする「動画データアップロードサービス」[13] が既に稼働している。このサービス上に録画ファイル群を登録すれば、ホストアカウントの利用者が自由にファイルを操作することができる。

### 5.1 NextCloud へのファイル登録

クラウドレコーディングからの一時ダウンロード先のディレクトリ構造(表 5)を読み込み、NextCloud 上のデータ保存エリアに登録するスクリプトを開発した。本スクリプトは el.kyutech.ac.jp ドメインに所

属するホストアカウント (xxxx.edu@el.kyutech.ac.jp) を対象としているが、他のドメインについても希望があれば移動可能となっている。NextCloud 上のデータ保存エリアには、ストリーミング用ディレクトリがあらかじめ用意 (表 6 の /aaaaaa) されているため、その中に、ホストアカウント名 (xxxx.edu) に対象年 (例 .2022) を結合したディレクトリが存在しなければ作成する。

図 8(上) にスクリプトが生成した実行シェルスクリプトの例を、図 8(下) に検証ログの例を示す。移動先のディレクトリが既に存在している場合は、mkdir は省略される。図 8(上) のスクリプトでは、mkdir の後に chown -R 48:48 を実行し、NextCloud が処理できるパーミッションに変更する。その後、cp -r を実行して複写を実行する。なお、複写元 (一時ダウンロード先) は別のタイミングで削除される。

表 5: 一時ダウンロード先のディレクトリ構造

	ドメイン名	年	ホスト名	ミーティング ID +レコーディング開始日時	ファイル群
例 1	/el.kyutech.ac.jp	/2022	/xxxx.edu	/.....	/.....
例 2	/jimu.kyutech.ac.jp	/2020	/yyyy	/.....	/.....

表 6: データ保存エリア先のディレクトリ構造

	九工大 ID	files	ストリーミング用ディレクトリ	ホスト名+年	ファイル群
例 1	/xxxx	/files	/aaaaaa	/xxxx.edu.2022	/.....
例 2	/zzzz	/files	/aaaaaa	/yyyy.2020	/.....

本シェルスクリプトの実行後に、sudo -u apache php occ files:scan --all のようなコマンドを実行し、NextCloud に対象ファイル群を認識させる。なお、全ユーザ (2024 年 3 月時点 17406 ユーザ) の再スキャンには 24 分程度の時間を要した。

```
#1,0,xxxx,xxxx.edu,mkdir:0
/bin/mkdir /nextcloud/xxxx/files/aaaaaa/xxxx.edu.2022
/bin/chown -R 48:48 /nextcloud/xxxx/files/aaaaaa/xxxx.edu.2022
/bin/cp -r /zoomdata/el.kyutech.ac.jp/2022/xxxx.edu/yyyyyy==2022.02.10.12.36UTC
  /nextcloud/xxxx/files/aaaaaa/xxxx.edu.2022/
/bin/chown -R 48:48 /nextcloud/xxxx/files/aaaaaa/xxxx.edu.2022
#1,1,xxxx,xxxx.edu,mkdir:SKIP
/bin/cp -r /zoomdata/el.kyutech.ac.jp/2022/xxxx.edu/zzzzzz==2022.02.10.07.07UTC
  /nextcloud/xxxx/files/aaaaaa/xxxx.edu.2022/
/bin/chown -R 48:48 /nextcloud/xxxx/files/aaaaaa/xxxx.edu.2022
```

```
Mkdir Nextcloud home area /nextcloud/xxxx/files/aaaaaa/xxxx.edu.2022/ (folder 1 meeting 0)
OK,1,0,2022,/zoomdata/el.kyutech.ac.jp/2022/xxxx.edu/yyyyyy==2022.02.10.12.36UTC,
  xxxx.edu,xxxx,yyyyyy==,2022.02.10.12.36UTC
OK,1,1,2022,/zoomdata/el.kyutech.ac.jp/2022/xxxx.edu/zzzzzz==2022.02.10.07.07UTC,
  xxxx.edu,xxxx,zzzzzz==,2022.02.10.07.07UTC
```

図 8: データ保存エリアに登録するスクリプトが生成したシェルスクリプトの例 (上) ・ログ (下)

ホストアカウント (xxxx.edu@el.kyutech.ac.jp) の「動画データアップロードサービス (NextCloud)」上での確認例を図 9 に示す。2021 年の保存データのため、ストリーミング用ディレクトリ (aaaaaa) の下



## 6 おわりに

ビデオ会議サービス (Zoom, WebEX) を用いた遠隔授業の経験は、多くの教職員にその活用方法や教育方法を浸透／定着させてきた。原則が対面授業となった現在においても、学習支援サービス (Moodle) を用いた教育手法は、2019 年度以前と比較すると増加しており、こちらも定着してきたと言える。

一方、フルオンラインの科目も現在でもいくつか開講されている。多くは、ビデオ会議サービス (Zoom) と学習支援サービス (Moodle) を組み合わせて遠隔授業を実施する。これにより、一般教養科目 (英語など) を担当することが多い非常勤講師が、遠隔授業を実施することで、大学への移動を減少させることができる。また、学習者もマイクを使うことができるため、発声させやすく、教員も発音を確認しやすいなどのメリットもある。その他にも、学習支援サービス (Moodle) の Reader プラグインを活用した英語多読の実践 [16] も進んできている点が注目すべき点である。今後も一般教養科目における遠隔授業をサポートしていく必要がある。

本稿では、ビデオ会議サービス (Zoom) を用いた教育／研究活動について、2020 年から 2022 年までの活用状況について解説した。2020 年 3 月に「ビデオ会議システムを用いた遠隔教育の実践」[17] として一度遠隔教育を取りまとめたが、ここ数年で新たな革新があったと感じている。また、本稿の後半では、クラウドレコーディングの残容量確保の手段として、Zoom API を活用した事例について説明した。本学では、Moodle API を用いた連携や NextCloud API といった API 連携の拡充に務めている。これは、このような API を用いた連携は、今後も重要な技術であり不可欠な手段であると考えているからである。

2023 年度からは、情報高度化本部を中心に、マイクロソフトの Teams サービスの提供が本格化し、キャンパス内電話網やビデオ会議サービスなどの活用が可能となった。これを受け、学習教育センターでも、ビデオ会議サービス (Teams) も教育・研究に活用できるように、準備を進めている。具体的には、学習教育センターが Teams 活用の講習会を開催している。なお、2024 年度の学習教育センター (教育高度化本部) は、新しいセンター長<sup>12</sup>の下、4 つのグループ (教育 DX 支援、教学 IR 支援、FD 支援、技術グループ) を再編して、より一層効果的な活動を行う予定である。

## 謝辞

本解説の成果の一部は、科学研究費補助金 (基盤研究 (C) JP20K03149, 基盤研究 (C) JP22K12297) の支援を受けた。なお、システム上における実践においては、大学改革推進等補助金 (デジタル活用教育高度化事業) 「学修活動分析を利用した教育高度化のためのデジタル活用仮想基盤整備」による支援を一部受けた。

## 参考文献

- [1] 山口真之介, 大西淑雅, 西野和典, 大橋 健, 篠原 武 : 免許法認定公開講座における同期型遠隔講義の実践, 九州工業大学情報科学センター広報, Vol. 1, No. 18, pp. 9–19 (2006).
- [2] 富重秀樹 : 九州工業大学における SCS 関連設備の整備について, 情報科学センター広報, Vol. 1, No. 18, pp. 81–90 (2006).
- [3] 大西淑雅, 山口真之介, 西野和典 : 九州工業大学における遠隔教育の実施支援, 大学 ICT 推進協議会 2018 年度年次大会, Vol. 2018, No. MB1-2, pp. 1–4 (2018).
- [4] 福田 豊, 畑瀬卓司, 富重秀樹, 林 豊洋 : BYOD 環境整備に向けた無線 LAN 通信実験, 情報処理学会論文誌, Vol. 60, No. 3, pp. 758–767 (2019).

<sup>12</sup><https://www.ltc.kyutech.ac.jp/center/dhead/> ※ 2023 年度までのセンター長は教育高度化本部長として関わる。

- [5] 林 豊洋, 大西淑雅, 山口真之介, 中山 仁, 福田豊他 3 名 : ノートパソコン必携化の支援を主眼とした教育研究用コンピュータシステムの更新, 情報処理学会研究報告, Vol. 2019-IOT-046, No. 13, pp. 1-7 (2019).
- [6] 日商エレクトロニクス : 導入事例紹介・九州工業大学, 日商エレクトロニクス (オンライン), <https://collab.nissho-ele.co.jp/case/kyutech-case.html> 2024-03-31.
- [7] 大西淑雅, 山口真之介 : 教育 IT 環境で利用可能な OSS 認証システムの活用, 大学 ICT 推進協議会 2021 年度年次大会, Vol. 2021, No. TD1-3, pp. 207-212 (2021).
- [8] 学習教育センター長 : ビデオ会議サービス (Zoom) の利用について (お知らせ), 九州工業大学 (オンライン), <https://www.ltc.kyutech.ac.jp/12551/> 2024-03-22.
- [9] Zoom 社 : Zoom オンプレミス導入, Zoom 社 (オンライン), [https://support.zoom.com/hc/ja/article?id=zm\\_kb&sysparm\\_article=KB0060087](https://support.zoom.com/hc/ja/article?id=zm_kb&sysparm_article=KB0060087) 2023-10-28.
- [10] aadistudio, A.: How can I get cloud Recordings video using zoom Api, Zoom Developers (online), <https://devforum.zoom.us/t/how-can-i-get-cloud-recordings-video-using-zoom-api/94828> 2024-03-31.
- [11] Rodrigues, R.: zoom-recording-downloader, Canada (online), <https://github.com/ricardorodrigues-ca> 2024-03-21.
- [12] SCIENCE, H.: The Moodle plugin, HUMAN SCIENCE (online), <https://hs-learning.jp/> 2024-02-21.
- [13] 学習教育センター長 : 動画データアップロードサービス, 九州工業大学 (オンライン), <https://www.ltc.kyutech.ac.jp/> 2024-04-08.
- [14] developers, Z.: Server-to-Server OAuth, Zoom 社 (online), <https://developers.zoom.us/docs/internal-apps/s2s-oauth/> 2024-03-31.
- [15] developers, Z.: Rate limits, Zoom 社 (online), <https://developers.zoom.us/docs/api/rest/rate-limits/> 2024-03-31.
- [16] 長瀬真理子 : 教養教育院言語系における特色ある取り組み「第 2 節 : 英語 XB : 多読 E ラーニングコース」, 教育ブレイク, Vol. 2022-2023, No. 9, pp. 97-100 (2024).
- [17] 山口真之介, 大西淑雅, 西野和典, 大橋 健, 篠原 武 : ビデオ会議システムを用いた遠隔教育の実践, THINK 会報第 118 号司法書士論叢, Vol. 1, No. 118, pp. 56-65 (2020).

◇◇◇◇◇  
解 説  
◇◇◇◇◇

## 九州工業大学における脆弱性の検査と改善の取り組み

佐藤 彰洋<sup>1</sup>  
福田 豊<sup>2</sup>  
中村 豊<sup>3</sup>

### 1 はじめに

昨今、国立大学法人において、サイバー攻撃によるセキュリティインシデントが多発している [1]。例えば、不正アクセスによる個人情報の漏洩やウェブサイトの改竄などの事案である。このようなセキュリティインシデントが発生した場合、法人としての信用失墜を招くだけでなく、その法人を取り巻く関係者に多大な影響を及ぼすことになる。故に、セキュリティインシデントに向けた体制と対策の整備は、法人全体が一丸となって取り組むべき責務となる。

九州工業大学では、「サイバーセキュリティ対策等基本計画」を策定し、情報セキュリティの向上に努めている。この基本計画で定められた一項目「情報機器の管理状況の把握及び必要な措置の実施」に則り、我々が属すネットワークセキュリティ基盤運用室では学外公開アドレスの厳格な管理に取り組んでいる [2]。ここで学外公開アドレスとは、学外から到達可能な IP アドレスを意味する。本稿では、学外公開アドレスの管理の一環として、そのアドレスを付与した機器に対する脆弱性検査の結果と、その改善状況の調査を中心に報告する。その脆弱性の検査と改善の結果、IP アドレスの学外公開が厳格に管理され、本学のネットワークが高い堅牢性を確保できることを確認した。

本稿の構成は次の通りである。先ず、2章で本学のネットワーク構成を説明する。次いで、学外公開アドレスを付与した機器に対する脆弱性検査の結果と、その改善状況の調査を3章で報告する。最後に、4章で本稿の貢献を纏める。

### 2 九州工業大学のネットワーク

九州工業大学におけるネットワーク構成を図1に示す [3]。本学が接続する SINET は、全国の教育研究機関の学術情報基盤として、国立情報学研究所が整備した情報通信ネットワークである。また2023年の時点では、学内外を分ける境界 FW システムとして米国 Fortinet 社の FortiGate 600E と米国 PaloAlto Networks 社の PA-5220 を直列に設置している [4, 5]。

本学では、我々が属すネットワークセキュリティ基盤運用室がコアネットワークの管理を、各部局がそれに接続する情報システムの管理を担当している。これは大学組織の業務が教育・研究・事務など多岐に渡るため、部局の意向を反映した情報システムの運用が不可欠となることに起因する。ここで特筆すべきは、ネットワークセキュリティ基盤運用室と各部局との一元的な対話のため、部局が情報システムごとに若干名の管理者（以降、情報システムの管理者と表記）を選任する点である。この情報システム

<sup>1</sup>情報基盤センター 助教 satoh@isc.kyutech.ac.jp

<sup>2</sup>情報基盤センター 准教授 fukuda@isc.kyutech.ac.jp

<sup>3</sup>情報基盤センター 教授 yutaka-n@isc.kyutech.ac.jp

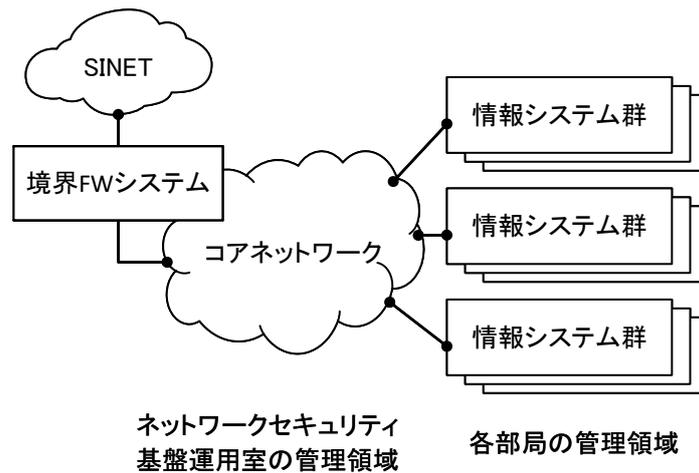


図 1: 九州工業大学のネットワーク構成

の独立性により、IP アドレスの学外公開は、情報システムの管理者からの申請に基づき、当該アドレスに対するサービス（ポートやプロトコル）単位の通信制御を適用している。なお、IP アドレスの学外公開に際して、情報システムの管理者には脆弱性検査とその改善を課している。この脆弱性検査には、米国 Tenable Network Security 社の Tenable Vulnerability Management を採用している [6]。これは Tenable Vulnerability Management が、ネットワークを介した通信のみから機器の脆弱性を検出する機能に加え、その脆弱性の深刻度（None, Low, Medium, High, Critical の 5 段階）および改善法を提示する機能を有するが故である。

### 3 脆弱性の検査と改善

#### 3.1 諸元

本学では、情報システムの管理者に対して、定期的な脆弱性の検査と改善を依頼している。ここで脆弱性の改善とは、深刻度 Medium 以上の脆弱性が学外に露呈することがないように対策を施すこととする。今回は、その脆弱性対策の期間を 2024 年 1 月から 2024 年 3 月末までの 3ヶ月間に設定した。次節では、その依頼の直前である 2023 年 12 月 1 日時点と、その期間中である 2024 年 3 月 15 日時点における脆弱性検査の結果とその改善状況の調査を中心に報告する。

#### 3.2 結果

脆弱性の検査と改善を依頼する直前である 2023 年 12 月 1 日時点で、学外公開アドレスの総数は 389 であった。表 1 と表 2 に、学外公開中の IP アドレスを付与した 389 台の機器に対する脆弱性検査の結果を示す。389 台の機器が有す脆弱性の総数は、Low が 238, Medium が 783, High が 884, Critical が 97 であった。それらは Low が 10 種類, Medium が 47 種類, High が 18 種類, Critical が 8 種類となっており、同一の脆弱性が多数検出されていることを確認した。脆弱性には偏りがあり、複数の機器間で同一の脆弱性が多数検出される傾向にあることを確認した。また、各機器において最も高い深刻度は、42 台が Low, 108 台が Medium, 5 台が High, 7 台が Critical を有しており、脆弱性が全く無い機器は 209 台であった。このことから、短い期間でも新たな脆弱性が発見されていることが見て取れる。

表 1: 機器が有す脆弱性の総数 (2023 年 12 月 1 日時点)

	Critical	High	Medium	Low	合計
総数	97	884	783	238	2002
種類	8	18	47	10	83

表 2: 脆弱性の深刻度と機器の数 (2023 年 12 月 1 日時点)

Critical	High	Medium	Low	None	合計
20	6	100	65	198	389

表 3: 機器が有す脆弱性の総数 (2024 年 3 月 15 日時点)

	Critical	High	Medium	Low	合計
総数	1	3	138	127	269
種類	1	1	16	10	28

表 4: 脆弱性の深刻度と機器の数 (2024 年 3 月 15 日時点)

Critical	High	Medium	Low	None	合計
1	3	62	40	153	259

脆弱性対策の期間中である 2024 年 3 月 15 日時点で、学外公開アドレスの総数は 389 であった。表 3 と表 4 に、そのアドレスを付与した 389 台の内、脆弱性の改善を終えた 259 台の結果を示す。259 台の機器が有す脆弱性の総数は、Low が 127、Medium が 138、High が 3、Critical が 1 であった。それらは Low が 10 種類、Medium が 16 種類、High が 1 種類、Critical が 1 種類となっており、特に Low と Medium において同一の脆弱性が多数検出されていることを確認した。また、各機器において最も高い深刻度は、40 台が Low、62 台が Medium、3 台が High、1 台が Critical を有しており、脆弱性が全く無い機器は 153 台であった。その結果における代表的な脆弱性の数と詳細を表 5 に示す。Critical の“PHP Unsupported Version Detection”に加え High の“CGI Generic SQL Injection (blind)”<sup>1</sup>は、誤検出であることを確認している。Medium は、SSL の自己証明書に起因するものが合計 67 件<sup>2</sup>、Git のリポジトリ公開に起因するものが合計 2 件<sup>3</sup>、VPN の共有鍵に起因するものが 1 件あり、サービスを停止する他に適当な手段が無いことから対処が不要の脆弱性と判断した。また、Medium の“HSTS Missing From HTTPS Server (RFC 6797)”の改善は、HTTP と HTTPS で同一コンテンツを表示することが求められるため、これも対処が不要の脆弱性とした。ここで“SSH Terrapin Prefix Truncation Weakness (CVE-2023-48795)”など、残りの Medium の脆弱性は、それに対する学外からの通信を遮断しているため、学外に露呈しているのは Low の脆弱性のみであることを補足しておく。故に、脆弱性の検査と改善により、ネットワークの堅牢性を低下させる要因を除外できたと言える。

<sup>1</sup>検査結果に“Note that this script is experimental and may be prone to false positives.”の記載がある

<sup>2</sup>表 5 の“SSL Certificate Cannot Be Trusted”, “SSL Self-Signed Certificate”, “SSL Certificate Expiry”が該当する

<sup>3</sup>表 5 の“Backup Files Disclosure”, “Git Repository Served by Web Server”が該当する

表 5: 代表的な脆弱性の数と詳細 (2024 年 3 月 15 日時点)

Critical	PHP Unsupported Version Detection	1
High	CGI Generic SQL Injection (blind)	3
Medium	SSL Certificate Cannot Be Trusted	37
Medium	SSL Self-Signed Certificate	24
Medium	SSL Certificate Expiry	6
Medium	HSTS Missing From HTTPS Server (RFC 6797)	36
Medium	SSH Terrapin Prefix Truncation Weakness (CVE-2023-48795)	19
Medium	Internet Key Exchange (IKE) Aggressive Mode with Pre-Shared Key	1
Medium	Backup Files Disclosure	1
Medium	Git Repository Served by Web Server	1
Low	SSH Weak Key Exchange Algorithms Enabled	53
Low	SSH Server CBC Mode Ciphers Enabled	25
Low	SSL Anonymous Cipher Suites Supported	18
Low	Web Server Allows Password Auto-Completion	14
Low	Web Server Uses Basic Authentication Without HTTPS	6
Low	SSH Weak MAC Algorithms Enabled	4

## 4 おわりに

本稿では、学外公開アドレスの管理の一環として、そのアドレスを付与した機器に対する脆弱性検査の結果と、その改善状況の調査を中心に報告した。その脆弱性の検査と改善の結果、IP アドレスの学外公開が厳格に管理され、本学のネットワークが高い堅牢性を確保できることを確認した。最後に、各情報システムの管理者の協力のもと、本稿で記述した脆弱性は既に改善されていることを特筆しておく。

**謝辞** 本研究は JSPS 科研費 JP21K11848 の助成を受けたものである。また、各情報システムの管理者には、本システムの運用にあたり多大な協力を頂いた。ここに深く謝意を示す。

## 参考文献

- [1] Trend Micro: 被害事例とリサーチから見る教育機関を狙うサイバー攻撃の動向, [https://www.trendmicro.com/ja\\_jp/jp-security/23/e/securitytrend-20230502-01.html](https://www.trendmicro.com/ja_jp/jp-security/23/e/securitytrend-20230502-01.html) (2023 年 7 月 10 日参照).
- [2] 佐藤彰洋 他: 学外公開アドレス管理システムの設計と評価, 情報処理学会デジタルプラクティス, Vol. 11, No. 3, pp. 624-635 (2020).
- [3] 中村豊 他: 九州工業大学における全学セキュア・ネットワークの更新, 情報処理学会研究報告, Vol. 2020-IOT-48, No. 28, pp. 1-6 (2020).
- [4] Fortinet: Global Leader of Cybersecurity Solutions and Services, <https://www.fortinet.com> (2023 年 7 月 10 日参照).

- [5] Palo Alto Networks: Leader in Cybersecurity Protection & Software for the Modern Enterprises, <https://www.paloaltonetworks.com> (2023年7月10日参照).
- [6] Tenable Network Security: Comprehensive Cybersecurity and Exposure Management, <https://www.tenable.com> (2023年7月10日参照).



◇◇◇◇◇  
解説  
◇◇◇◇◇

## DGA マルウェアにより生成された悪性ドメインの検出

佐藤 彰洋<sup>1</sup>  
福田 豊<sup>2</sup>  
中村 豊<sup>3</sup>

### 1 はじめに

マルウェアはインターネットにおける重大な脅威のひとつである。サイバー犯罪者は、C&C (Command-and-Control Server) を介してマルウェアに感染した端末を操作することで、機密情報窃取、フィッシング詐欺、標的型攻撃などの悪意ある活動を試みる。米 McAfee 社の報告によると、約 30 万のマルウェアが日々誕生しており、それによる世界の総損失額は年間 6000 億ドルを超える [1]。そのため、マルウェアに対抗するための技術の確立が急務である。

マルウェアによる被害の抑止のため、管理者は自身のネットワークに内在する感染端末を迅速に排除することが求められる。一方、多くのマルウェアには、検出を回避するための機能として DGA (Domain Generation Algorithm) が実装されている [2]。DGA とは、C&C のドメインを頻繁に変更することで、マルウェアから C&C へ向けた通信であるコールバックを隠蔽するための仕組みである。具体的には、マルウェアは DGA に基づいて多数のドメインを自動生成した後、それらドメインの名前解決を試みる。その名前解決の結果、正しい応答を返したドメインを C&C のものと見做し、そのドメインとの間で通信を確立する。ここで留意すべきは、コールバックのために生成されるドメインの生存時間が極端に短い点である。故に、従来のブラックリストを用いた通信の監視では、DGA マルウェアのコールバックを検出することが困難となる [3]。

マルウェアと C&C との通信を補足するために、パケットのペイロードを参照する DPI (Deep Packet Inspection) が用いられてきた [4, 5]。一方、2017 年の時点でインターネットにおける暗号化通信の割合は 50% を超えること、それと併せて約 70% のマルウェアが通信を暗号化することが確認されている [6]。このように暗号化は情報保護の一般的な手段であり、それが占める割合に反比例して DPI を適用可能な通信は極僅かなもののみとなっている。

本稿では、DNS (Domain Name System) に対する膨大な数の名前解決から、DGA マルウェアのコールバックのために自動生成したドメインの判別を試みる。DNS に着目した理由は、マルウェアによる通信に先んじて必ず名前解決が生じること、その名前解決は暗号化による通信内容の隠蔽が困難であることに起因する。DGA マルウェアにより生成された悪性ドメインは生存時間が極端に短いため、その判別に利用可能な特徴が限定的である。その問題を踏まえ、我々は表層的な文字列解析に基づく悪性ドメイン判別手法を提案する。本手法の独自性は、DGA に関する事前情報を全く必要とせず、ドメイン文字列の意味の有無からドメインの良性と悪性を推定する点にある。これは、人為的に生成された良性ドメインが、組織や商品の名称など、その意図を反映した文字列を成すこと、自動的に生成された悪性ドメインが、登録済みのドメインとの衝突を避けるため無意味な文字列を成すことから、良性ドメインと

<sup>1</sup>情報基盤センター 助教 satoh@isc.kyutech.ac.jp

<sup>2</sup>情報基盤センター 准教授 fukuda@isc.kyutech.ac.jp

<sup>3</sup>情報基盤センター 教授 yutaka-n@isc.kyutech.ac.jp

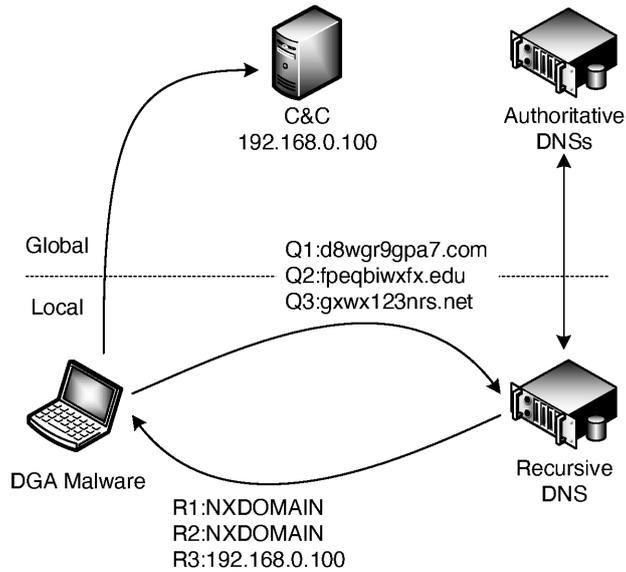


図 1: DGA マルウェアと C&C とのコールバック通信

悪性ドメインの文字列において明確な差異が現れるが故である。また実験を通じて、提案手法が 0.9960 の再現率と 0.9029 の適合率で悪性ドメインを判別可能であることを確認した。すなわち、悪性ドメインの名前解決を予兆として DGA マルウェアを高精度で検出できることを示した。この結果から、ネットワークに内在するマルウェアへの迅速な対処が可能となるため、ネットワークにおける安全性の向上が期待できる。

本稿の構成は次の通りである。まず、2 章で既存研究とその問題点を整理する。3 章で文字列解析の基づく悪性ドメイン判別手法を提案した後、4 章で提案手法の有効性を議論する。最後に 5 章で本研究の貢献と課題を纏める。

## 2 関連技術

本章では、DGA マルウェアを中心とした関連技術について述べる。2.1 節で DGA マルウェアの詳細について説明した後、2.2 節で既存研究とその問題点を整理する。

### 2.1 DGA マルウェア

Conficker や Kraken など、世界で深刻な被害を齎したマルウェアは、その機能の一部として DGA が実装されている。また、それらマルウェアを広範囲に拡散するため、ウェブページやウェブ広告への不正コードの埋め込みが観測されている [7, 8]。

図 1 に DGA マルウェアによるコールバックの概要を示す。ここで、図中の **Q** で示す通信はマルウェアから RDNS (Recursive DNS Server) <sup>1</sup> に対する名前解決を、**R** はその応答を意味する。また、C&C のものとして、DGA により生成されたドメインが事前に ADNS (Authoritative DNS Server) に登録されているものとする。まず、マルウェアは DGA に基づいて複数のドメインを自動生成し、それらドメインを自身の属するネットワーク内の RDNS に問い合わせる。RDNS は、ドメインが登録済みであった場

<sup>1</sup>RFC8499 で定義される Recursive Resolver を指す。Full-Service Resolver とは、キャッシュ機能の有無により区別される。

合、そのドメインに対応付けられたアドレスを、ドメインが未登録であった場合、エラーメッセージとして NXDOMAIN (Non-Existing Domain) を応答する。最終的に、マルウェアは正しい応答があったドメインを C&C のものと見做し、そのドメインに対してコールバックを試みる。

DGA の目的は、マルウェアと C&C の間に可用性の高い通信経路を確立することにある。具体的には、C&C のドメインを変更することで、ブラックリストに基づく通信の遮断を容易に回避することが可能となる。加えて、ネットワーク内から外へ向けた通信は宛先が多岐に渡るためコールバックの発見が困難となること、アドレス変換やファイアウォールにより通信を制限されないことが挙げられる。ここで留意すべきは、マルウェアと C&C とで同一 DGA を用いることで、ドメインの変更に一切の情報交換を必要としない点である。

## 2.2 既存研究と問題点

ブラックリストの高度化については頻繁な研究が行われており、現在もネットワークにおける脅威防御戦略の中核を成している。Soldo らは、複数の参加者から提供される過去の攻撃ログに基づいて、新たにブラックリストを生成する方法を提案している [9]。また、Freudiger らは、P2P の技術を応用することで、機密性を担保した攻撃ログの共有を実現している [10]。それに対して、DGA マルウェアは、C&C のドメインを頻繁に変更することにより、ブラックリストに基づく通信の遮断を回避する機能を有している。

Gu らは、DPI に基づく受動的なネットワーク監視システムとして BotHunter を実装した [11]。BotHunter は、マルウェアの一般的な挙動をモデル化して、それと関連の強い通信を感染の根拠とする。また、DPI の性能改善に向けた取り組みなどが報告されている [12, 13]。しかしながら、ネットワークにおける暗号化の担う役割に反比例して、DPI を適用可能な通信は極僅かなもののみとなっている。故に、マルウェアの検出のための情報源として、暗号化の影響を受けない DNS の名前解決が注目されている。

Rahbarinia らは、DNS の名前解決において既知の悪性ドメインと高確率で共起するドメインから未知の悪性ドメインを発見する Segugio を開発した [14]。Segugio は次の直感的知見、(1) 同一マルウェアファミリーに感染した端末は、同一悪性ドメイン群と通信する傾向にあること、(2) 未感染の端末は、悪性ドメインと通信することがないことに基づいている。一方、DGA マルウェアにおいては、コールバック通信に生存時間が極端に短い一時的な悪性ドメインを用いるため、その一時的な悪性ドメインと共起するドメインは存在し得ない。故に、このシステムは DGA マルウェアの通信に対して効果を成し得ない。

Bilge らは、DNS の名前解決とその応答から計測可能な特徴量と機械学習を用いてドメインを評価する Exposure を開発した [15]。特徴量の例は、ドメインの生存期間、ドメインに割り当てられたアドレスの数、ドメインの文字長などである。このシステムが採用する教師有り機械学習において、その精度は一般的に学習用データセットの数と質に依存する。しかしながら、DGA マルウェアにおける多くの悪性ドメインは NXDOMAIN を応答するため特徴量が得られないこと、C&C に対応付けられた悪性ドメインは生存時間が極端に短いことから、十分な量の学習用データセットを確保することが困難となる。

Berger らは、名前解決におけるアドレスとドメインの関係の変化を継続的に学習する DNSMap を構築した [16]。DNSMap は、C&C のアドレスが複数のドメインに、そのドメインが複数のアドレスに対応付けられること、それらの対応関係が時間経過に伴い急速な変化を示すことに着眼している。一方、Wang らは、名前解決の挙動と分布特性に基づいて DGA マルウェアの検出する DBod を実装した [17]。その検出は、同一 DGA により生成された候補ドメイン群に対して接続を試みるため、同一マルウェアファミリーに感染した端末による名前解決が特定の期間中に高い類似性を示すことに基づいている。これらのシステムは広範囲に渡る DNS トラフィックの観測を必要とするため、その適用は ISP などの大規模なネットワークに限定される。

これまでに、マルウェアが生成するドメインの解析結果や [18], その解析を自動化する仕組みなどが報告されている [19]. その結果を踏まえ、幾つかの研究では文字列の特徴のみを用いたドメインの判別が試みられている. Truong らは、ドメイン文字列のみから良性と悪性を判別する手法を提案した [20]. この手法は、教師有り機械学習とバイグラムモデリングによりドメインにおける頻出文字パターンを学習する. Anderson らは、深層学習を用いた文字レベルのモデリングにより、その手法を拡張した [21]. また、Vinayakumar らにより、多様な機械学習と深層学習を用いた判別精度の比較結果が示されている [22]. これらの手法は、悪性ドメインを生成するためのルールに識別可能な偏りが存在することに基づいている. しかしながら、その識別には事前の学習を必要とするため、未知の DGA マルウェアに対する性能の低下が懸念される.

### 3 提案

本稿では、DNS に対する膨大な数の名前解決から、DGA マルウェアにより自動生成されたドメインの判別を試みる. 表 1 に、DGA マルウェアにより生成されたドメインの例を示す. 2.2 節で述べた通り、同一マルウェアファミリーにおける名前解決の類似性を利用するためには、広範囲に渡る DNS トラフィックの観測が必要となる. それに加え、DGA マルウェアにおける悪性ドメインは生存時間が極端に短いため、その判別に利用可能な特徴が限定的である. その問題を踏まえ、我々は表層的な文字列解析に基づく悪性ドメイン判別手法を提案する. その理由は、(1) 人為的に生成された良性ドメインが、組織や商品の名称など、その意図を反映した文字列を成すこと、(2) 自動的に生成された悪性ドメインが、登録済みのドメインとの衝突を避けるため意味の無いランダムな文字列を成すこと、(3) それ故に、良性ドメインと悪性ドメインの文字列において明確な差異が現れることに起因する. マルウェアの検出を容易にするためのランダム化の概念は新しいものではなく、この概念に基づく幾つかの研究が発表されている [20, 21, 22, 23, 24]. 例えば、Lin らはサイバー犯罪者へのトレースバックのためにマルウェアの通信を解読する手法を [23], Wahab らはマルウェアに侵害された仮想マシンを検出するための手法を提案している [24]. また、ドメイン文字列の特徴のみを用いた良性と悪性の判別が試みられている [20, 21]. なお、文献 [20] と [21] の両手法については、提案手法との類似性から 4 章にて定量的な判別性能の比較を実施した. ここで特筆すべきは、本手法は DGA に関する事前情報を全く必要とせず、ドメイン文字列の意味の有無からドメインの良性と悪性を推定する点にある.

図 2 に提案手法の概要を示す. 本手法は 3 つの機能、(1) サブドメイン選択機能、(2) 辞書に基づく判別

表 1: DGA マルウェアにより生成されたドメインの例

ChinAd	xe0d7fazyrvvw19f.ru 7qvdqaw561dtasyi.com
Conficker	xjjjvqpoh.com.ai pfnnwjoeuee.biz
Locky	bt1wubflhf11shn.info jlbroeji.biz
NewGOZ	lygx14u1vnf8hb1twhv8619h8ygr.net cipu0wdgsnq9u8st8m1lym0hq.com
Nymaim	embonxn.info ghhimbgrpx.biz

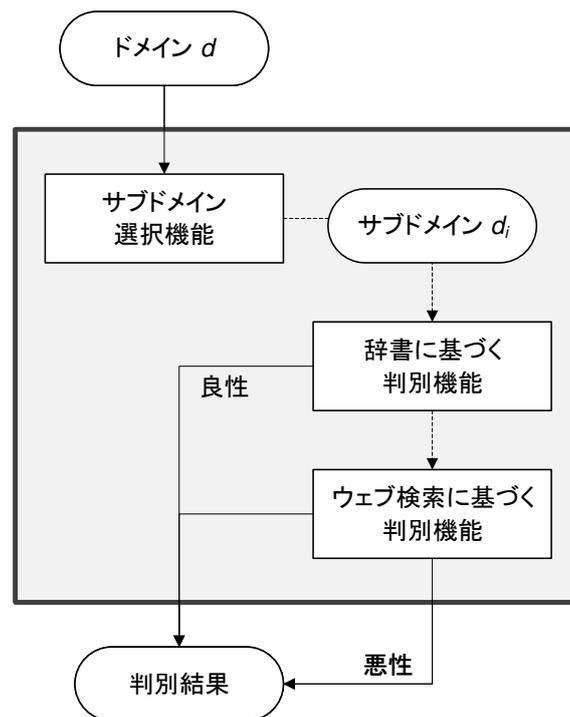


図 2: 表層的な文字列解析に基づく悪性ドメイン判別手法の概要

機能, (3) ウェブ検索に基づく判別機能により構成される。以降の節で、各機能の詳細について述べる。

### 3.1 サブドメイン選択機能

本機能は、効率的な判別のため、ドメイン  $d$  から文字長が最大のサブドメイン  $d_i$  を選択する。ここで、 $d_i$  はドメイン  $d$  を構成する  $i$  レベルのサブドメインを意味する。例えば、ドメイン  $d$  が `xjjjvqpoh.com.ai` である場合、文字長が最大のサブドメインは  $d_3$  の `xjjjvqpoh` となる。これは、短い文字列のドメインは正当な組織に占有されている可能性が高いこと、登録済みのドメインとの衝突を避ける必要があることから、DGA はドメインとして比較的長い文字列を生成することに起因する。

### 3.2 辞書に基づく判別機能

本機能は、まず辞書  $\mathbb{D}$  を参照することにより、サブドメイン  $d_i$  の文字列を単語群  $w$  に分割する。次いで、その単語群  $w$  の特徴からサブドメイン  $d_i$  のランダム性を推定する。その結果、サブドメイン  $d_i$  が意味の有る文字列であると判断された場合、それを含むドメイン  $d$  を人為的に生成された良性ドメインと見做す。

ドメイン文字列の単語分割において参照するため、6種類の辞書を準備した。そのひとつは、ウェブクロウリングで作成したコーパスと英語辞書から成る。他の5つは、上述の辞書にフランス語、ドイツ語、スペイン語、ロシア語、および日本語の辞書を追加したものである。それら各言語の辞書は、ドメインの表記を考慮して非アルファベットをアルファベットで置換している。その置換では、発音記号を取り除き、例えば  $\acute{a}$  を  $a$  で、 $\hat{i}$  を  $i$  で、 $\ddot{u}$  を  $u$  で、 $\csc$  を  $c$  で表記した。また、ドイツ語の  $\text{ß}$  を  $ss$  で、スペイン語の  $\tilde{n}$  を  $n$  と  $nn$  で、日本語はヘボン式ローマ字で、ロシア語のキリル文字は文献 [25] の規則に基づいて置き換えた。なお、スペイン語の  $\tilde{n}$  のように、一般的に複数の代用表記が用いられる場合は、そ

れら規則を1単語に対して適用している. 具体的には, スペイン語 año の場合, ano と anno の2単語が辞書に登録されることになる.

式1により, (1) 各単語の文字長が最大, 且つ単語数が最小になること, (2) 極端な選択率の差により辞書に含まれる単語を優先することの, 2つの観点に基づいてサブドメイン  $d_i$  の文字列を単語群  $w$  に分割する.

$$\mathcal{F}(d_i) = \arg \max_{w \in \mathbb{W}(d_i)} \frac{1}{n} \prod_{j=1}^n \mathcal{P}(w_j)$$

$$\mathcal{P}(w_j) = \begin{cases} 1 & (w_j \in \mathbb{D}) \\ 1/|\mathbb{D}|^{|w_j|} & (w_j \notin \mathbb{D}) \end{cases} \quad (1)$$

ここで,  $\mathbb{W}(d_i)$  はサブドメイン  $d_i$  の文字列における全分割候補の集合,  $w$  は単語  $w_1, \dots, w_j, \dots, w_n$  から成る候補の単語群,  $|w_j|$  は単語  $w_j$  の文字長,  $|\mathbb{D}|$  は辞書  $\mathbb{D}$  の総単語数をそれぞれ意味する. また,  $\mathcal{P}(w_j)$  は, 単語  $w_j$  が辞書  $\mathbb{D}$  に含まれるか否かに基づいて, 単語  $w_j$  の選択率を導出する関数である. この処理において, 6種類の辞書を参照した単語分割から最良の結果を選択する. その結果,  $n=1$ , すなわちサブドメイン  $d_i$  自体が辞書  $\mathbb{D}$  に含まれる場合, ドメイン  $d$  を良性ドメインと判断する.

サブドメイン  $d_i$  が人為的に生成された文字列である場合, 単語群  $w$  において各単語の文字長が長いこと, 且つ単語数が少ないことが特徴として挙げられる. その特徴を踏まえ, 式2によりサブドメイン  $d_i$  のランダム性を推定する.

$$y_\alpha = \sum_{k=1}^n u_k \mathcal{L}_k(w) \quad (2)$$

ここで  $u_k$  は,  $\mathcal{L}_k(w)$  に重みを付与する係数を, また  $\mathcal{L}_k(w)$  は, 単語群  $w$  を文字長に基づいて降順に並び替え, その  $k$  番目の値を導出する関数を意味する. 例えば, 単語群  $w$  が kyutech, local, domain, name から成る場合,  $\mathcal{L}_2(w)$  は domain の文字長である6を,  $\mathcal{L}_4(w)$  は name の文字長である4を出力する. 特筆すべきは, 辞書に含まれない単語は, その文字長を0と見做す点である. この結果,  $y_\alpha > th_\alpha$  を満たす場合, ドメイン  $d$  を良性ドメインと判断する.

### 3.3 ウェブ検索に基づく判別機能

本機能は, 辞書の不足を補うため, ウェブ検索の結果を参考にサブドメイン  $d_i$  のランダム性を推定する. 例えば, 辞書を持たない言語のドメイン, 固有名詞を用いたドメイン, 頭字語を用いたドメイン, 国際化ドメインなどは辞書に基づく文字列解析で良性と悪性を判別できない.

まず, (1) サブドメイン  $d_i$  の文字列との完全一致を検索すること, (2)  $th_\beta$  番目以降の候補を表示することの, 2つの条件を満たすクエリをウェブ検索エンジンに対して発行する. これは, 文字列における意味の有無が, それに関連付けられたウェブページの数から間接的に推定可能であることに着眼している. 式3にウェブ検索エンジンに対するクエリの例を示す.

$$str = "d_i" \& \text{num} = th_\beta \quad (3)$$

ここで,  $str$  は検索対象の文字列を指定する変数であり, その二重引用符により検索結果を文字列  $d_i$  と完全一致するウェブページのみ限定している. また,  $num = th_\beta$  は, 検索結果の表示を  $th_\beta$  番目以降のウェブページとすることを意味する. 留意すべきは, xn-11q68wkwbj6u や xn-sjqw6xkwbgyd9ay881 など, 文字列  $d_i$  が xn-- で始まる国際化ドメインである場合, それを事前に元の文字列に復号する点である. その結果,  $y_\beta > th_\beta$ , すなわち文字列  $d_i$  が関連付けられたウェブページの数  $y_\beta$  が閾値  $th_\beta$  を超

える場合、ドメイン  $d$  を人為的に生成された良性ドメインと判断する。それ以外を DGA により自動的に生成された悪性ドメインと見做す。

表 2 に、サブドメインのウェブ検索結果を示す。表から明らかのように、これは各ドメインの良悪を反映した指標となっている。

## 4 評価

本節では、実験を通じた提案手法の評価により、DNS に対する膨大な数の名前解決から悪性ドメインを判別できること、その結果に基づくことで DGA マルウェアのコールバックを高精度で検出できることを示す。4.1 節で実験の諸元について述べた後、4.2 節と 4.3 節で結果について議論する。

### 4.1 諸元

提案手法との比較のため、文献 [20] と [21] を参考にドメインの文字列のみから悪性ドメインを判別する 2 種類の方法を実装した。それらの実装に用いた機器の構成を表 3 に示す。第一の実装は、ドメイン文字列に対するバイグラムモデリングと教師有り機械学習により判別する手法であり、第二の実装は、

表 2: サブドメインのウェブ検索結果

Benign Domains	Web Search Result
kyutech	175,000
comsoc	133,000
centrum24	66,400
incometaxindiaefiling	117,000
curapelanatureza	791,000
xn-sjqw6xkwbgyd9ay88l	789,000
Malicious Domains	Web Search Result
7qvdqaw561dtasyi	0
pfnnwjoeuee	1
jlbroeji	1
cipu0wdgsnq9u8st8m1lym0hq	0
ghhimbgbbrpx	0
ixajqhvaegcqgrhiwyv	1

表 3: 実装に用いた機器の構成

CPU	Xeon Silver 4110 (8core 2.10GHz)
GPU	NVIDIA GeForce RTX 2080 Ti (11GB GDDR6)
RAM	96GB DDR4-2666
SSD	Seq. Read and Write up to 560MB/sec and 530MB/sec
Kernel	Linux 3.10.0-957.1.3.el7.x86_64
Software	TensorFlow 2.0.0, CUDA Toolkit 10.0, cuDNN 7.6.5

表 4: 各データセットにおけるドメインの数

<i>D0</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>	<i>D5</i>	<i>D6</i>	<i>D7</i>	<i>D8</i>	<i>D9</i>
Alexa 1000000	ChinAd 3000	Conficker 3000	CoreBot 3000	Fobber 3000	Kraken 3000	Locky 3000	NewGOZ 3000	Nymaim 3000	PadCrypt 3000
<i>D10</i>	<i>D11</i>	<i>D12</i>	<i>D13</i>	<i>D14</i>	<i>D15</i>	<i>D16</i>	<i>D17</i>	<i>D18</i>	<i>D19</i>
Prosliekfan 3000	Pykspa 3000	Qadars 3000	Ramnit 3000	Shiotob 3000	Simda 3000	Symmi 3000	Tempedreve 3000	Tinba 3000	Vawtrak 3000

表 5: 実験結果

	Recall	Precision
Truong et al. [20]	0.7915	0.5366
Anderson et al. [21]	0.8447	0.6462
Our work	0.9960	0.9029

深層学習のひとつである LSTM (Long Short Term Memory) ネットワークを用いた文字レベルのモデリングにより判別する手法である。ここで留意すべきは、提案手法とは異なり、これらの実装は判別のために良性と悪性ドメインのデータセットによる学習が必須な点である。

表 4 に実験に用いたデータセットを示す。悪性ドメインは、逆行解析を通じて明らかになったマルウェアの DGA により生成したものである [26]。実験に用いたマルウェアは、ChinAd, Conficker, Locky, NewGOZ, Nymaim などの計 19 種である。また、良性ドメインは Alexa [27] が公開するアクセス数の上位 1,000,000 である。2 種類の実装のために、良性ドメインから 5%、全悪性ドメインから 15% を無作為に抽出したものを学習用データセットとした。残りが検証用データセットである。ネットワークにおいて実測されたデータに代わり、これらデータセットを採用した理由は、良性と悪性が正確に付与されたドメインでない厳密な精度の比較が困難なことに起因する。

提案手法における辞書  $\mathbb{D}$  として、ウェブクロールで作成したコーパス [28, 29]、および Aspell [30] に登録されている単語を使用した。その総単語数は約 8,000,000 語である。また、各種パラメタを経験的に  $u_1 = 2$ ,  $u_2 = 1$ ,  $u_3 = 0.25$ ,  $th_\alpha = 15$ ,  $th_\beta = 50$  に設定した。その  $y_\beta$  は、bing.com と yahoo.com の 2 つのウェブ検索エンジンによる結果の内、大きい方の値とした。これらの最適化は今後の課題とする。

## 4.2 定量的評価

各手法における悪意ドメインの判別性能を定量的に評価するために、一般的な 2 つの指標を用いた。再現率は、悪性ドメインの総数に対する悪性と判別されたドメインの数の比率であり、適合率は、悪性と判別されたドメインの総数に対する真に悪性であるドメインの数の比率である。

実験結果を表 5 に示す。この結果から、提案手法は 0.9960 の再現率と 0.9029 の適合率を達成しており、2 つの実装よりも高い精度を示すことが見て取れる。2 つの実装の精度が低下した理由は次の通りである。まず、学習用データの量が不十分であったこと、学習用データに偏りがあったことが挙げられる。次いで、単純な文字の並びのみからドメインの良性と悪性を判別することの限界である。提案手法は、データセットを用いた事前の学習が不要であること、ドメインの文字列解析により良性と悪性を判

表 6: 各データセットにおけるドメインの誤判数

	$D_0$	$D_1$	$D_2$	$D_3$	$D_4$	$D_5$	$D_6$	$D_7$	$D_8$	$D_9$
Truong et al. [20]	33108	110	660	98	340	432	509	18	652	282
Anderson et al. [21]	22407	22	522	34	146	202	274	2	514	222
Our work	5189	0	102	0	0	2	1	0	25	0
	$D_{10}$	$D_{11}$	$D_{12}$	$D_{13}$	$D_{14}$	$D_{15}$	$D_{16}$	$D_{17}$	$D_{18}$	$D_{19}$
Truong et al. [20]	674	702	278	448	246	638	1586	632	314	1480
Anderson et al. [21]	576	580	286	218	150	626	1272	396	152	1326
Our work	19	21	0	0	0	3	1	0	0	17

別することから、精度低下の要因を除外できたと考えられる。

各データセットにおいて誤判したドメインの数を表6に示す。提案手法は、14種類のマルウェアにおける誤判は5ドメイン未満であり、非常に高い精度を達成している。しかしながら、Nymaim, Proslifean, Pyksa, および Vawtrak は誤判は20ドメイン程度、AlexaとConfickerに至っては誤数が5189と102ドメインとなり、他と比較して精度の低下が目につく結果となった。Confickerが生成するドメインを調査したところ、長さが4から12までのランダムな文字列であった。これらの中で誤判されたのは、5文字以下の文字列から成るドメインのみである。すなわち、ドメインの文字長が短い場合、頭字語を用いた良性ドメインと機械的に生成された悪性ドメインを区別できないことが誤判の原因である。その他のマルウェアにおけるドメインの誤判は、長さ6前後の文字列が偶然にも意味の有る単語を成したことに起因している。その具体例は、docket.com, wouled.biz, olleman.comである。一方、Alexaにおける誤判は次の4種、(1) 頭字語を用いたドメイン、(2) 非アルファベットをアルファベット表記したドメイン、(3) 数字を多く含むドメイン、(4) ランダムな文字列から成るドメインであった。ここで、(1)のドメインはConfickerと同様の原因、(2)と(3)のドメインは、その辞書を持たないことが原因である。(4)のドメインは、良性と悪性の差が文字列に現れないため、提案手法による判別が仕組み上困難である。その具体例は、31qjnuhra3xf585jgkxhk71exuhu6yrkna.com や ctxxgxdnhctxxgx dnh.xyz である。

提案手法における辞書に基づく判別機能のパラメタ  $u_k$  と  $th_\alpha$  は、ドメインを良性と判別する基準となる、文字列  $d_i$  に占める単語の割合を決めるものである。本実験を通じて、この機能により Alexa の 541029 ドメインを正確に良性と判別できたこと、それに対して悪性を良性と誤判したのが Nymaim の 11 ドメインと Pyksa の 7 ドメインのみであったことから、それらパラメタは概ね適当な値であったと言える。一方、ウェブ検索に基づく判別機能のパラメタ  $th_\beta$  は、ウェブ検索結果に強く依存した値となっている。そのため、判別の精度に強く影響を及ぼすのは、パラメタの微細な設定よりもウェブ検索エンジンの選択であると考えられる。具体的には、Alexaに含まれる fatosdesconhecidos.com.br は、yahoo.comによる検索では  $th_\beta$  の値を下回り悪性と誤判されるが、yahoo.brによる検索では正確に良性と判別されることを確認している。このことから、特に非アルファベットをアルファベット表記したドメインの場合、その ccTLD (Country Code Top Level Domain) に応じたウェブ検索エンジンの結果を参照することで更なる精度の向上が期待できる。

以上の議論より、幾つかの課題があるにしても提案手法が 0.9960 の再現率と 0.9029 の適合率で悪性ドメインを判別できることを確認した。この結果は、これら悪性ドメインの名前解決を予兆として、ネットワークに内在する DGA マルウェアを高精度で検出できることを示唆している。

表 7: 提案手法と既存手法との定性的な比較

	(1) DGA Malware Detection	(2) On-Line Detection	(3) Robust to Encryption	(4) Network- Scale Independent	(5) No Need for a Priori Dataset
Soldo et al. [9]	Poor	Good	Good	Good	Poor
Gu et al. [11]	Poor	Good	Poor	Good	Good
Rahbarinia et al. [14]	Poor	Good	Good	Fair	Poor
Bilge et al. [15]	Poor	Fair	Good	Good	Poor
Berger et al. [16]	Good	Good	Good	Poor	Fair
Wang et al. [17]	Good	Good	Good	Poor	Good
Truong et al. [20]	Fair	Good	Good	Good	Poor
Anderson et al. [21]	Fair	Good	Good	Good	Poor
Our work	Good	Poor	Good	Good	Good

### 4.3 定性的評価

表 7 に提案手法と既存手法との定性的な比較を示す。その比較の観点は、(1) DGA マルウェアの検出性能、(2) 検出の実時間性、(3) 暗号化に対する頑健性、(4) ネットワークの規模に対する依存の有無、(5) データセットを用いた学習など事前知識の有無である。先ず前節で述べたように、事前の学習を必要とすることなく、提案手法は DGA マルウェアの高精度な検出を実現している。機械学習や深層学習に基づく手法は [20, 21]、判別精度の維持に潤沢な学習用データセットの準備が必須となるが、そのデータセットに対して良性と悪性のラベルを付与する作業は非常に煩雑である [31]。加えて、Sivaguru らは、最新の研究成果の比較を通じて、ホワイトリストやブラックリストを学習用データセットとして使用することに実用性の観点から疑問を呈している [32]。これは、ホワイトリストやブラックリストが、ネットワークで実測される良性ドメインと悪性ドメインの特徴を十分に反映するものでないことに起因する。それに対して、本手法が頼るのは DGA マルウェアとは直接関係しない辞書とウェブ検索のみである。これらは一般に公開されているものであり、その利用が容易である点を留意されたい。

提案手法による検出は、Gu らの BotHunter [11] とは異なり、マルウェアによる通信の暗号化の影響を受けないこと、Berger らの DNSMap [16] や Wang らの DBod [17] とは異なり、大規模なネットワークの観測を必要としないことが挙げられる。特に、ネットワークにおける暗号化通信の重要性を踏まえ、これまでに DoT (DNS over TLS) の標準化が推進されている。このプロトコルは、DNS の名前解決を TLS でカプセル化することにより、盗聴や改竄などを防ぐことを目的としている。一方、提案手法が良性と悪性を判別するために必要とするのは、ドメインの文字列のみである。従って、ネットワーク内の RDNS が記録するシステムログの情報のみで動作可能であり、DoT による暗号化の影響を受けることは無い。但し、RDNS に対する名前解決が生じない DoH (DNS over HTTPS) については、本手法を適用することはできない。

提案手法の欠点は、ウェブ検索の結果に頼るため、他と比較して良性と悪性の判別に時間を要する点にある。各機能における 1 秒当りに処理可能なドメイン数を算出すると、サブドメイン選択機能と辞書に基づく判別機能はそれぞれ 2968.51 と 1341.39 であるのに対して、ウェブ検索に基づく判別機能は 1.35 と非常に低速であるが故に、ネットワーク内で生じる全ての名前解決に適用することは困難である。加えて、判別するドメイン数の増加に伴い、ウェブ検索エンジンに対して多大な負荷をかけることとなる。一方、文献 [33] では、DGA によるコールバック先の変更に伴い NXDOMAIN が発生することに着

目している。この知見を踏まえ、NXDOMAIN の頻度から感染の疑われる端末を絞り込むこと、その幾つかの名前解決に対して本手法を適用することで、それら問題の大幅な緩和が期待できる。

Sood らは、自身が有す辞書の単語を連結することで悪性ドメインを生成する DGA マルウェアの代表例として、Rovnix について言及している [34]。その様な悪性ドメインの文字列は良性ドメインのものと類似した特徴を有すため、提案手法による判別は困難である。具体的には、`accelerateaccountant.in.net` や `accelerateactress.in.net` など、Rovnix が生成したドメインは、提案手法における辞書に基づく判別機能により良性と見做される。この問題の解決には、Pereira らの取り組みのように、ドメイン文字列における単語間の関係性を考慮するなどの対策が必要となる [35]。

特筆すべきは、非常に限定された情報のみに依存するため、提案手法は高い汎用性を有す点である。具体的には、学習用データセットなどの事前知識を必要とせず、その判別はドメイン文字列のみに基づいている。故に、他手法の機能の一部として組み込むことが容易であり、その判別性能の改善に大きく寄与すると考えられる。

## 5 おわりに

本稿では、DGA マルウェアの検出のため、表層的な文字列解析に基づく悪性ドメイン判別手法を提案した。その独自性は、DGA に関する事前情報を全く必要とせず、ドメイン文字列の意味の有無からドメインの良性と悪性を推定する点にある。また実験を通じて、提案手法が 0.9960 の再現率と 0.9029 の適合率で悪性ドメインを判別可能であること、すなわち悪性ドメインの名前解決を予兆として DGA マルウェアを高精度で検出できることを確認した。この結果から、ネットワークに内在するマルウェアへの迅速な対処が可能となるため、ネットワークにおける安全性の向上が期待できる。

特筆すべき特徴は、本手法が有す高い汎用性である。これは、高精度な判別を実現することに加え、その判別はドメイン文字列のみに基づくこと、学習用データセットなどの事前知識を必要としないことに起因する。故に、他手法の機能の一部として組み込むことが容易であり、その判別性能の改善に大きく寄与すると考えられる。

今後は、大規模なネットワークで観測した通信を対象に、本手法の有効性を評価する予定である。また、他手法との組み合わせの検討と、それが判別性能に与える効果を明らかにする。

**謝辞** 本研究は JSPS 科研費 JP21K11848 の助成を受けたものである。ここに深く謝意を示す。

## 参考文献

- [1] J. A. Lewis: Economic Impact of Cybercrime — No Slowing Down, <https://www.csis.org/analysis/economic-impact-cybercrime> (2018).
- [2] Y. Fu et al.: Stealthy Domain Generation Algorithms, *IEEE Transactions on Information Forensics and Security*, Vol. 12, No. 6, pp. 1430–1443 (2017).
- [3] M. Kühner et al.: Paint It Black: Evaluating the Effectiveness of Malware Blacklists, *Proceedings of the International Symposium on Research in Attacks, Intrusions and Defenses*, pp. 1–21 (2014).
- [4] Z. Chen et al.: Malware Characteristics and Threats on the Internet Ecosystem, *Journal of Systems and Software*, Vol. 85, No. 7, pp. 1650–1672 (2012).
- [5] T. Nelms et al.: ExecScent: Mining for New C&C Domains in Live Networks with Adaptive Control Protocol Templates, *Proceedings of the USENIX Conference on Security Symposium*, pp. 589–604 (2013).

- [6] Cisco Systems Inc.: Cisco Annual Cybersecurity Report, [https://www.cisco.com/c/dam/m/hu\\_hu/campaigns/security-hub/pdf/acr-2018.pdf](https://www.cisco.com/c/dam/m/hu_hu/campaigns/security-hub/pdf/acr-2018.pdf) (2018).
- [7] D. Kim: Potential Risk Analysis Method for Malware Distribution Networks, *IEEE Access*, Vol. 7, pp. 185157–185167 (2019).
- [8] C. Dwyer et al.: Malvertising — A Rising Threat to the Online Ecosystem, *Journal of Information Systems Applied Research*, Vol. 10, No. 3, pp. 29–37 (2017).
- [9] F. Soldo et al.: Blacklisting Recommendation System: Using Spatio-Temporal Patterns to Predict Future Attacks, *IEEE Journal on Selected Areas in Communications*, Vol. 29, No. 7, pp. 1423–1437 (2011).
- [10] J. Freudiger et al.: Controlled Data Sharing for Collaborative Predictive Blacklisting, *Proceedings of International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, pp. 327–349 (2015).
- [11] G. Gu et al.: BotHunter: Detecting Malware Infection Through IDS-Driven Dialog Correlation, *Proceedings of the USENIX Conference on Security Symposium*, pp. 167–182 (2007).
- [12] N. Cascarano al.: Optimizing Deep Packet Inspection for High-Speed Traffic Analysis, *Journal of Network and Systems Management*, Vol. 19, No. 1, pp. 7–31 (2011).
- [13] T. J. Parvat al.: Performance Improvement of Deep Packet Inspection for Intrusion Detection, *Proceedings of the IEEE Global Conference on Wireless Computing & Networking*, pp. 224–228 (2014).
- [14] B. Rahbarinia et al.: Efficient and Accurate Behavior-Based Tracking of Malware-Control Domains in Large ISP Networks, *ACM Transactions on Privacy and Security*, Vol. 19, No. 2, pp. 4:1–4:31 (2016).
- [15] L. Bilge et al.: Exposure: A Passive DNS Analysis Service to Detect and Report Malicious Domains, *ACM Transactions on Information and System Security*, Vol. 16, No. 4, pp. 14:1–14:28 (2014).
- [16] A. Berger et al.: Mining Agile DNS Traffic Using Graph Analysis for Cybercrime Detection, *Computer Networks*, Vol. 100, pp. 28–44 (2016).
- [17] T. S. Wang et al.: DBod: Clustering and Detecting DGA-based Botnets using DNS Traffic Analysis, *Computers & Security*, Vol. 64, pp. 1–15 (2017).
- [18] D. Plohmann et al.: A Comprehensive Measurement Study of Domain Generating Malware, *Proceedings of the USENIX Conference on Security Symposium*, pp. 263–278 (2016).
- [19] G. Vighianisi et al.: SysTaint: Assisting Reversing of Malicious Network Communications, *Proceedings of the Software Security, Protection, and Reverse Engineering Workshop*, pp. 1–12 (2018).
- [20] D. Truong et al.: Detecting Domain-Flux Botnet based on DNS Traffic Features in Managed Network, *Security and Communication Networks*, Vol. 9, No. 14, pp. 2338–2347 (2016).
- [21] H. S. Anderson et al.: DeepDGA: Adversarially-Tuned Domain Generation and Detection, *Proceedings of the ACM Workshop on Artificial Intelligence and Security*, pp. 13–21 (2016).
- [22] R. Vinayakumar et al.: Evaluating Deep Learning Approaches to Characterize and Classify the DGAs at Scale, *Journal of Intelligent and Fuzzy Systems*, Vol. 34, No. 3, pp. 1265–1276 (2018).

- [23] W. Lin et al.: Traceback Attacks in Cloud – Pebbletrace Botnet, *Proceedings of International Conference on Distributed Computing Systems Workshops*, pp. 417–426 (2012).
- [24] O. A. Wahab et al.: Optimal Load Distribution for the Detection of VM-based DDoS Attacks in the Cloud, *IEEE Transactions on Services Computing*, Vol. 13, No. 1, pp. 114–129 (2020).
- [25] LinguaJunkie.com: Russian Alphabet Guide, [https://www.linguajunkie.com/wp-content/uploads/RussianAlphabetGuide\\_1\\_12-1.pdf](https://www.linguajunkie.com/wp-content/uploads/RussianAlphabetGuide_1_12-1.pdf).
- [26] J. Bader: Some Results of My DGA Reversing Efforts, [https://github.com/baderj/domain\\_generation\\_algorithms](https://github.com/baderj/domain_generation_algorithms).
- [27] Hacker Target Pty. Ltd.: Download Top 1 Million Sites, <https://hackertarget.com/top-million-site-list-download/>.
- [28] P. Norvig: Natural Language Corpus Data: Beautiful Data, <http://norvig.com/ngrams/>.
- [29] Linguatools: Wikipedia Monolingual Corpora, <https://linguatools.org/tools/corpora/wikipedia-monolingual-corpora/>.
- [30] K. Atkinson: GNU Aspell, <http://aspell.net>.
- [31] Y. Roh et al.: A Survey on Data Collection for Machine Learning: A Big Data — AI Integration Perspective, *IEEE Transactions on Knowledge and Data Engineering*, Vol. PP, No. 99, pp. 1–1 (2019).
- [32] R. Sivaguru et al.: An Evaluation of DGA Classifiers, *Proceedings of the IEEE International Conference on Big Data*, pp. 5058–5067 (2018).
- [33] M. Antonakakis et al.: From Throw-Away Traffic to Bots: Detecting the Rise of DGA-Based Malware, *Proceedings of the USENIX Conference on Security Symposium*, pp. 491–506 (2012).
- [34] A. K. Sood et al.: A Taxonomy of Domain-Generation Algorithms, *IEEE Security & Privacy*, Vol. 14, No. 4, pp. 46–53 (2016).
- [35] M. Pereira et al.: Dictionary Extraction and Detection of Algorithmically Generated Domain Names in Passive DNS Traffic, *Proceedings of the International Symposium on Research in Attacks, Intrusions, and Defenses*, pp. 295–314 (2018).



◇◇◇◇◇  
解説  
◇◇◇◇◇

## 辞書に基づく DGA マルウェアにより生成された悪性ドメインの検出

佐藤 彰洋<sup>1</sup>  
福田 豊<sup>2</sup>  
中村 豊<sup>3</sup>

### 1 はじめに

マルウェアはインターネットにおける重大な脅威のひとつである。サイバー犯罪者は、C&C (Command-and-Control Server) を介してマルウェアに感染した端末を操作することで、機密情報窃取、フィッシング詐欺、標的型攻撃などの悪意ある活動を試みる。米 McAfee 社の報告によると、亜種を含めて約 30 万のマルウェアが日々誕生しており、それによる世界の総損失額は年間 6000 億ドルを超える [1]。そのため、マルウェアに抗する技術の確立が急務である。

マルウェアによる被害の抑止のため、管理者は自身のネットワークに内在する感染端末を迅速に排除することが求められる。一方、多くのマルウェアには、検出を回避するための機能として DGA (Domain Generation Algorithm) が実装されている [2]。DGA とは、C&C のドメインを頻繁に変更することで、マルウェアから C&C へ向けた通信であるコールバックを隠蔽するための仕組みである。具体的には、マルウェアは DGA に基づいて多数のドメインを機械的に生成した後、それらドメインに対して名前解決を試みる。その結果、正しい応答を返したドメインを C&C のものと見做し、そのドメインとの間で通信を確立する。

幾つかの研究では、良性と悪性のドメイン文字列の差異から DGA マルウェアのコールバック通信を検出している [3, 4, 5]。これは、登録済みのドメインとの衝突を避けるため、悪性ドメインが無意味な文字列から成ることに起因する。それに対して、これまでの検出を無効化する高度な DGA マルウェアが出現している。このマルウェアは、それ自身が有する辞書の単語を連結することで、人為的なものと判別が困難なドメインを機械的に生成する [6]。

本稿では、DNS (Domain Name System) に対する膨大な数の名前解決要求から、辞書に基づく DGA マルウェアにより生成されたドメインの判別を試みる。DNS に着目した理由は、マルウェアによる通信に先んじて必ず名前解決が生じること、暗号化による通信内容の隠蔽が困難であることに起因する。我々は、良性ドメインと悪性ドメインでは文字列で頻出する単語や共起する単語に明確な差異が現れるという仮定を踏まえ、ドメインの文字列を構成する単語の関係性に基づく悪性ドメイン判別手法を提案する。これにより、文字列のみからドメインの良性と悪性の判別が可能となる。また実験を通じて、提案手法が 0.9977 の再現率と 0.9869 の適合率で悪性ドメインを判別可能であることを確認した。すなわち、悪性ドメインの名前解決を予兆として辞書に基づく DGA マルウェアを高精度で検出できることを示した。この結果から、ネットワークに内在する多様なマルウェアへの迅速な対処が可能となるため、ネットワークの運用において安全性の向上が期待できる。

---

<sup>1</sup>情報基盤センター 助教 satoh@isc.kyutech.ac.jp

<sup>2</sup>情報基盤センター 准教授 fukuda@isc.kyutech.ac.jp

<sup>3</sup>情報基盤センター 教授 yutaka-n@isc.kyutech.ac.jp

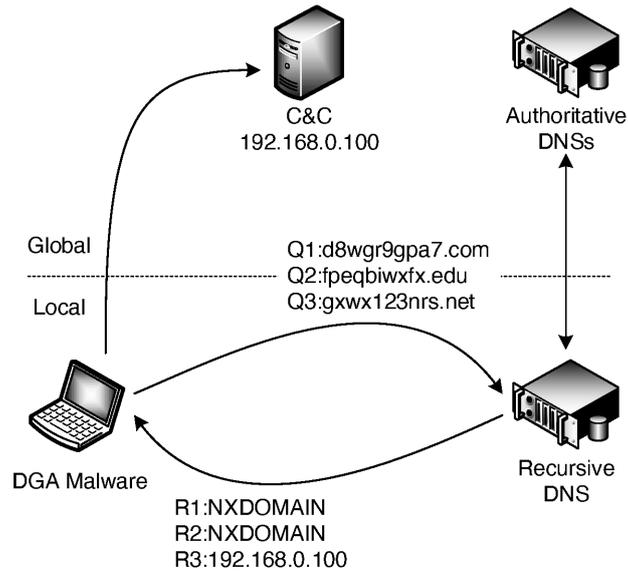


図 1: DGA マルウェアと C&C とのコールバック通信

本稿の構成は次の通りである。まず、2章で既存研究とその問題点を整理する。3章でドメイン文字列を構成する単語の関係性に基づく悪性ドメイン判別手法を提案した後、4章で提案手法の有効性を議論する。最後に5章で本研究の貢献と課題を纏める。

## 2 関連技術

本章では、マルウェアを中心とした関連技術について述べる。2.1節でDGAマルウェアの詳細について説明した後、2.2節で既存研究とその問題点を整理する。

### 2.1 DGA マルウェア

Conficker や GameOver Zeus, Torpig など、世界で深刻な被害を齎したマルウェアには、その機能の一部としてDGAが実装されている。また、それらマルウェアを広範囲に拡散するため、ウェブページやウェブ広告への不正コードの埋め込みが観測されている [7, 8]。

図1にDGAマルウェアによるコールバックの概要を示す。ここで、図中のQで示す通信はマルウェアからRDNS (Recursive DNS Server)<sup>1</sup>に対する名前解決を、Rはその応答を意味する。また、C&Cのものとして、DGAにより生成されたドメインが事前にADNS (Authoritative DNS Server)に登録されているものとする。まず、マルウェアはDGAに基づいて複数のドメインを機械的に生成し、それらドメインを自身の属するネットワーク内のRDNSに問い合わせる。RDNSは、ドメインが登録済みであった場合、そのドメインに対応付けられたアドレスを、ドメインが未登録であった場合、エラーメッセージとしてNXDOMAIN (Non-Existing Domain)をそれぞれ応答する。最終的に、マルウェアは正しい応答があったドメインをC&Cのものと同見做し、そのドメインに対してコールバックを試みる。

良性ドメインと比較して、一般的なDGAマルウェアによるドメインは文字列に明確な差異が見取れる。例えば、1ygx14u1vnf8hb1twhv8619h8ygr.net や cipu0wdgsnq9u8st8m1lym0hq.com

<sup>1</sup>RFC8499で定義されるRecursive Resolverを指す。Full-Service Resolverとは、キャッシュ機能の有無により区別される。

など、GameOver Zeus のドメインは文字長が 15 から 30 までのランダムな英数字から成る [9]. これは、既に登録済みのドメインとの衝突を避けることが狙いである. 一方、辞書に基づく DGA マルウェアは、自身が有する辞書の単語を連結することで、`accelerateaccountant.in.net` や `accelerateactress.in.net` など、人為的に生成したものと判別が困難なドメインを機械的に生成する [6].

DGA の目的は、マルウェアと C&C の間に可用性の高い通信経路を確立することにある. 具体的には、C&C のドメインを変更することで、ブラックリストに基づく通信の遮断を容易に回避することが可能となる. 加えて、ネットワーク内から外へ向けた通信は宛先が多岐に渡るためコールバックの発見が困難となること、アドレス変換やファイアウォールにより通信を制限されないことが挙げられる. ここで留意すべきは、マルウェアと C&C とで同一 DGA を用いることで、ドメインの変更に一切の情報交換を必要としない点である.

## 2.2 既存研究と問題点

ブラックリストの高度化については継続的な研究が行われており、現在もネットワークにおける脅威防御戦略の中核を成している. Soldo らは、複数の参加者から提供される過去の攻撃ログに基づいて、ブラックリストを更新する方法を提案している [10]. また、Freudiger らは、P2P の技術を応用することで、機密性を担保した攻撃ログの共有を実現している [11]. それに対して、DGA マルウェアは、C&C のドメインを頻繁に変更することにより、ブラックリストに基づく通信の遮断を回避する機能を有している.

マルウェアと C&C との通信を補足するために、パケットのペイロードを参照する DPI (Deep Packet Inspection) が用いられてきた. Gu らは、DPI に基づく受動的なネットワーク監視システムとして BotHunter を実装した [12]. BotHunter は、マルウェアの一般的な挙動をモデル化して、それと関連の強い通信を感染の根拠とする. また、DPI の性能改善に向けた取り組みなどが報告されている [13]. 一方、2017 年の時点でインターネットにおける暗号化通信の割合は 50% を超えること、それと併せて約 70% のマルウェアが通信を暗号化することが確認されている [14]. この暗号化の普及に反比例して、DPI を適用可能な通信は極僅かなもののみとなっている. 故に、マルウェアの検出のための情報源として、暗号化の影響を受けない DNS の名前解決が注目されている.

Rahbarinia らは、DNS の名前解決において既知の悪性ドメインと高確率で共起するドメインから未知の悪性ドメインを発見する Segugio を開発した [15]. Segugio は次の直感的知見、(1) 同一マルウェアファミリーに感染した端末は、同一悪性ドメイン群と通信する傾向にあること、(2) 未感染の端末は、悪性ドメインと通信することがないことに基づいている. 一方、DGA マルウェアにおいては、コールバック通信に生存時間が極端に短い一時的な悪性ドメインを用いるため、その一時的な悪性ドメインと共起するドメインは存在しない. 故に、このシステムは DGA マルウェアの通信に対して効果を成し得ない.

Berger らは、名前解決におけるアドレスとドメインの関係の変化を継続的に学習する DNSMap を構築した [16]. DNSMap は、C&C のアドレスが複数のドメインに、そのドメインが複数のアドレスに対応付けられること、それらの対応関係が時間経過に伴い急速な変化を示すことに着眼している. 一方、Wang らは、名前解決の挙動と分布特性に基づいて DGA マルウェアの検出する DBod を実装した [17]. その検出は、同一 DGA により生成された候補ドメイン群に対して接続を試みるため、同一マルウェアファミリーに感染した端末による名前解決が特定の期間中に高い類似性を示すことに基づいている. これらのシステムは広範囲に渡る DNS トラフィックの観測を必要とするため、その適用は ISP (Internet Service Provider) などの大規模なネットワークに限定される.

これまでに、DGA マルウェアが生成するドメインについての報告がなされている [18]. それを踏まえ、文字列の特徴のみを用いたドメインの判別が試みられてきた. Truong らは、ドメイン文字列のみから良性・悪性を判別する手法を提案した [4]. この手法は、教師有り機械学習とバイグラムモデリングによりドメインにおける頻出文字パターンを学習する. Anderson らは、深層学習を用いた文字レベ

ルのモデリングにより、その手法を拡張した [5]. 加えて、Vinayakumar らにより、多様な機械学習と深層学習を用いた判別精度の比較が示されている [19]. これらの手法は、悪性ドメインを生成するためのルールに識別可能な偏りが存在することに基づいている. しかしながら、単語を考慮しない文字レベルのモデリングでは、辞書に基づく DGA マルウェアの検出において十分な精度が期待できない.

本稿と同様に、辞書に基づく DGA マルウェアに焦点を当てたものとして Pereira らの取組みがある [20]. この手法は、DGA マルウェアの有する辞書を、それが生成したドメインから再構築することに主眼を置いている. 一方、良性と悪性の判別は非常に単純で、その辞書にドメイン文字列を成す単語が閾値以上含まれるか否かに基づいている. 故に、多岐に渡る DGA に対して精度を維持するためには、その判別の仕組みの高度化が必須である.

### 3 提案

本稿では、DNS に対する膨大な数の名前解決から、辞書に基づく DGA マルウェアにより生成されたドメインの判別を試みる. 表 1 に、辞書に基づく DGA マルウェアが生成したドメインの例を示す. これらドメインはコールバック先の C&C に依存するため、その生成には各マルウェアで異なる辞書とアルゴリズムが用いられることとなる. 例えば、文献 [21] と [22] で報告されている Gozi と Matsnu の辞書を比較したところ、単語数は 975 と 1,391 であり、それら辞書間の重複は 240 のみであった<sup>2</sup>. それに加え、マルウェアが機械的に生成したドメインは特定の辞書の単語から構成されるため、人為的に生成したドメインと比較して使用される単語に大きな偏りが生じると予想される. そのドメイン文字列で頻出する単語や共起する単語に明確な差異が現れるという仮定を踏まえ、我々はドメインの文字列を構成する単語の関係性に基づく悪性ドメイン判別手法を提案する. 本手法の特徴は、(1) 文字列のみからドメインの良性と悪性を判別すること、(2) ドメイン文字列を成す単語群の関係を一般的なグラフ理論における重み付き無向グラフで表現すること、(3) グラフにおける各頂点の中心性により各単語の重要性を測ることにある. 最終的に、その指標に基づく特徴ベクトルに機械学習アルゴリズムを適用することで、ドメインの良性と悪性を判別する.

図 2 に提案手法の概要を示す. 本手法は、(1) ノイズ除去機能、(2) 単語分割機能、(3) 単語グラフ構築機能、(4) 特徴ベクトルに基づく学習・分類機能により構成される. 3.1 節で学習用データセットについて、それ以降の節で各機能の詳細について述べる. なお、本手法では、2.1 節で述べた DGA マルウェアにおけるコールバック通信の特徴を踏まえ、良性と悪性を判別するドメインの数を大きく絞り込んでいる. 具体的には、DNS の名前解決要求がドメイン名からアドレスへの変換であること、その結果として

表 1: 辞書に基づく DGA マルウェアと悪性ドメインの例

Banjori	earnestnessbiophysicalohax.com pbmnestnessbiophysicalohax.com
Gozi	williamseasily.com printingthatlabel.com
Matsnu	shoulderracerecognizeblue.com emergencyadaptselectdoubt.com
Suppobox	windowtherefore.net severadifference.net

<sup>2</sup>文献 [21] で記載のある Ursnif は Gozi の別称である.

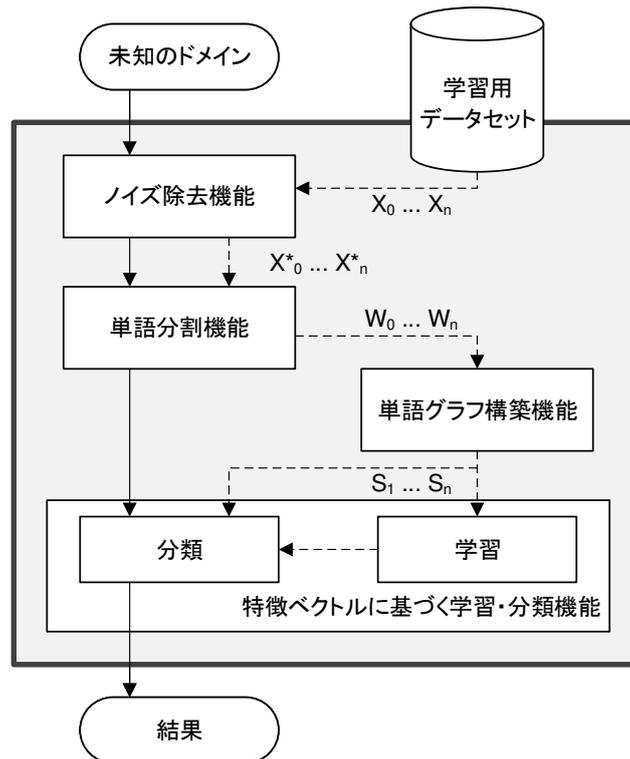


図 2: ドメインの文字列を構成する単語間の関係性に基づく悪性ドメイン判別手法の概要

NXDOMAIN を応答することの、両条件を満たすもののみを未知のドメインと学習用データセットとして用いる。

### 3.1 学習用データセット

提案手法において、特定のドメインのみを含む数種のデータセット  $X_i, i \in 0 \dots n$  を学習のために用いることとなる。ここで、 $X_0$  に属するのは良性ドメイン、 $X_1$  から  $X_n$  に属するのは  $n$  種の DGA マルウェアに起因する悪性ドメインとする。特筆すべきは、各ドメインをプライマリドメインまで短縮する点である。プライマリドメインは登録可能な最高レベルのサブドメインである [23]。具体的には、`www.ipsj.or.jp` と `smtp.isc.kyutech.ac.jp` のプライマリドメインは、それぞれ `ipsj.or.jp` と `kyutech.ac.jp` となる。

### 3.2 ノイズ除去機能

本機能は、先ずホワイトリストと合致したものを自明な良性ドメインと判断する。ホワイトリストに含まれるのは、`kyutech.jp` や `kyutech.ac.jp` などの自組織が有するドメイン、`uribl.com` や `dnswl.org` などの DNSBL・DNSWL (DNS Blacklists and Whitelists) に代表される特定のサービスが利用するドメイン、`trendmicro.com` や `barracudacentral.org` などのセキュリティ製品に関連付けられたドメインである。それに加え、DNS の仕様違反する文字列から成るドメインを [24]、入力ミスや設定ミスに起因するノイズと見做す。なお、本手法は国際化ドメインに非対応であるため [25]、ここで併せて `xn--` を接頭辞とするドメインを除外する点を留意されたい。その残りのデータセット  $X_i^*, i \in 0 \dots n$  は、さらなる分析のために次の機能に渡される。

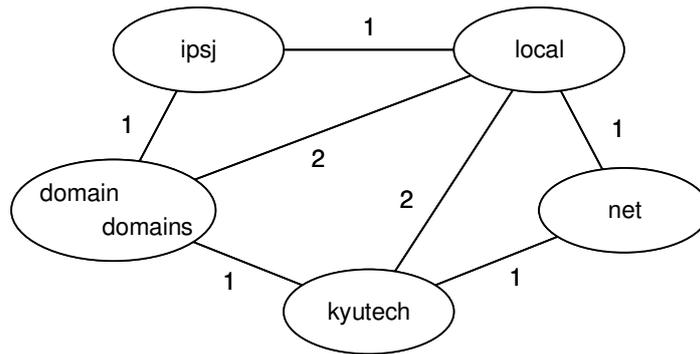


図 3: ドメイン文字列を成す単語群を用いた単語グラフの構築例

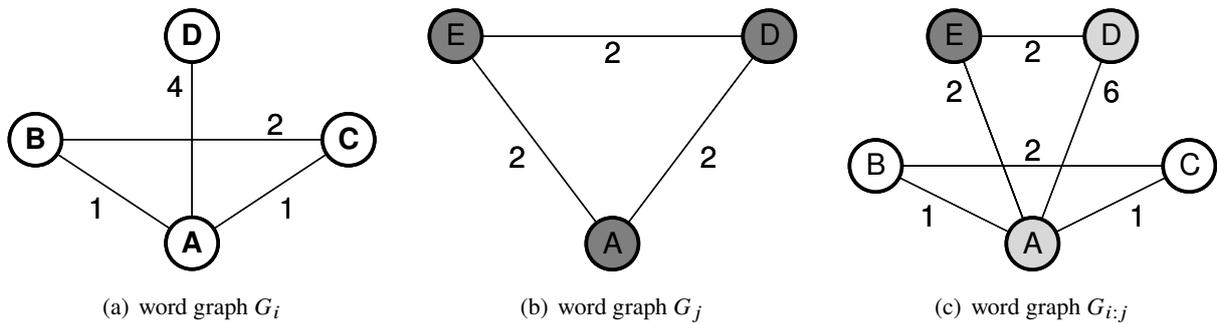


図 4: 単語グラフ  $G_i$  と  $G_j$  の結合とその結果  $G_{i,j}$

### 3.3 単語分割機能

本機能は、辞書  $\mathbb{D}$  に基づいてドメイン  $x$  のプライマリレベルの文字列を単語群  $w$  に分割する。辞書  $\mathbb{D}$  に含まれるのは、クローリングで作成したコーパスと英語辞書である。その分割を次式で示す。

$$\mathcal{F}(x) = \arg \max_{w \in \mathbb{W}(x)} \frac{1}{m} \prod_{j=1}^m \mathcal{P}(w_j)$$

$$\mathcal{P}(w_j) = \begin{cases} 1 & (w_j \in \mathbb{D}) \\ 1/|\mathbb{D}|^{|w_j|} & (w_j \notin \mathbb{D}) \end{cases}$$

ここで、 $\mathbb{W}(x)$  はドメイン  $x$  のプライマリレベルの文字列における全分割候補の集合、 $w$  は単語  $w_1, \dots, w_j, \dots, w_m$  から成る候補の単語群、 $|w_j|$  は単語  $w_j$  の文字長、 $|\mathbb{D}|$  は辞書  $\mathbb{D}$  の総単語数をそれぞれ意味する。また、 $\mathcal{P}(w_j)$  は、単語  $w_j$  が辞書  $\mathbb{D}$  に含まれるか否かに基づいて、単語  $w_j$  の選択率を導出する関数である。この結果は、(1) 各単語の文字長が最大且つ単語数が最小になること、(2) 極端な選択率の差により辞書に含まれる単語を優先することの、両条件を満たす分割となる。なお、文字列 `localdomain` の分割が `local` と `domain` の 2 単語から成る場合、その結果を  $\{\text{local}, \text{domain}\}$  と表記することとする。最終的に、本機能は学習用データセット  $X_i^*$  におけるドメインから、そのプライマリレベルの文字列を成す単語群の集合  $W_i$  を出力する。

### 3.4 単語グラフ構築機能

本機能は、単語グラフを用いることで、プライマリレベルの文字列を成す単語群の集合  $W_i$  の関係性を示す。単語グラフ  $G_i$  は、頂点と辺の集合から成る重み付き無向グラフである。頂点は集合  $W_i$  における単語、辺の重みは単一ドメインにおいて2つの単語が共起する頻度を意味する。ここで留意すべきは、単語の活用を考慮するため、文字列が類似した単語を同一頂点に集約する点である。その類似性の導出には編集距離比を、集約には閾値  $th$  の重心法に基づく階層型クラスタリングを採用した [26]。図3に、 $\{\text{kyutech, local, domain}\}$ ,  $\{\text{local, kyutech, net}\}$ ,  $\{\text{ipsj, domains, local}\}$ の単語群を用いた単語グラフの構築例を示す。この単語グラフにおいて、 $\text{domain}$ と $\text{domains}$ は文字列の類似性から同一頂点に集約されることになる。

次いで、単語グラフにおける処理としてグラフの結合を、指標として単語の重要性を定義する。結合は、単語グラフ  $G_i$  と  $G_j$  を構築するために用いた単語群の集合  $W_i$  と  $W_j$  の和から、単語グラフ  $G_{i,j}$  を再構築する処理である。その例を図4に示す。図4(a)と4(b)の単語グラフ  $G_i$  と  $G_j$  の結合は、図4(c)の単語グラフ  $G_{i,j}$  となる。なお、白色の頂点は単語グラフ  $G_i$  のみ含まれる単語、黒色の頂点は単語グラフ  $G_j$  のみ含まれる単語、灰色の頂点は両グラフに共通する単語であることを意味する。重要性は、単語グラフにおける中心性に関連付けられた指標である。任意の単語群  $w$  の重要性を次式で示す。

$$S_i(w) = \sum_{w_j \in w} |w_j| (C_{0:i}(w_j) - C_0(w_j))$$

ここで、 $|w_j|$  は単語  $w_j$  の文字長を意味する。また、 $C_0(w_j)$  と  $C_{0:i}(w_j)$  は、単語グラフ  $G_0$  と  $G_{0:i}$  における単語  $w_j$  の中心性を導出する関数である。機械的に生成した悪性ドメインと人為的に生成した良性ドメインでは、その文字列で頻出する単語や共起する単語に明確な差異が現れるが故に、良性と悪性の判別に効果的な単語は単語グラフにおいて中心的な役割を担うこととなる。その観点を踏まえ、良性データセットから構築した単語グラフ  $G_0$  を基準として、それと悪性データセットから構築した単語グラフ  $G_i$  の結合による中心性の変化量を単語  $w_j$  の重要性とした。その中心性の導出には、単語グラフが無向且つ非連結になることを勘案して PageRank を採用した [27]。

### 3.5 特徴ベクトルに基づく学習・分類機能

本機能は、先ず学習用データセットのドメインから特徴ベクトルを導出する。単語グラフにおける重要性に基づいて、単語群  $w$  から成るドメイン文字列の特徴ベクトルを次式で示す。

$$\vec{w} = (S_1(w), \dots, S_i(w), \dots, S_n(w))$$

次いで、それら特徴ベクトルに機械学習アルゴリズムを適用することで学習モデルを構築する。機械学習アルゴリズムには、その汎化性能と判別性能を勘案して SVM (Support Vector Machine) を採用した [28]。最終的に、未知のドメインはノイズの除去、ドメイン文字列の分割と特徴ベクトルの導出を経て、その学習モデルに基づくことにより良性と悪性が判別される。

## 4 評価

本節では、実験を通じた提案手法の評価により、DNS に対する膨大な数の名前解決から悪性ドメインを判別できること、その結果に基づくことで辞書に基づく DGA マルウェアのコールバックを高精度で検出できることを示す。4.1 節で実験の諸元について述べた後、4.2 節と 4.3 節で結果について議論する。

表 2: 各データセットにおけるドメインの数

$X_0$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$
Benign 3021124	Banjori 30000	Gozi 30000	Matsnu 30000	Pizd 30000	Rovnix 30000	Sisron 30000	Suppobox 30000

表 3: 実装に用いた機器の構成

CPU	Xeon Silver 4110 (8core 2.10GHz)
GPU	NVIDIA GeForce RTX 2080 Ti (11 GB GDDR6)
RAM	96 GB DDR4-2666
SSD	Seq. Read and Write up to 560 MB/sec and 530 MB/sec
Kernel	Linux 3.10.0-957.1.3.el7.x86_64
Software	TensorFlow 2.0.0, CUDA Toolkit 10.0, cuDNN 7.6.5

## 4.1 諸元

表 2 に実験に用いたデータセットを示す. 良性ドメイン  $X_0$  は, キャンパスネットワークの RDNS で観測された, 名前解決要求がドメイン名からアドレスへの変換であり, その結果として NXDOMAIN を応答したドメイン群である. キャンパスネットワークは計 6,000 人を超える学生と職員が利用しており, そのネットワークに接続する端末の名前解決を RDNS が担っている. 計測の期間は 2018 年 8 月の 1ヶ月間で, 条件を満たす応答の数は全名前解決要求の 1.5% にあたる 3,021,124 であった. それらドメイン群に悪性ドメインが含まれないことの調査として, 複数のセキュリティ製品により感染が疑われる端末を選定した後に, その端末が文献 [18, 6, 17] で報告のある “短期間に見慣れないドメインの名前解決と NXDOMAIN 応答が生じる” という DGA マルウェアの特徴と合致しないことを確認した. この調査で悪性ドメインの混在が無いことを完全に担保できるわけではない. しかしながら, 良性ドメインに混在する悪性ドメインの数は極微量となるため, 実験結果に対して大きな影響を及ぼさないと考えられる. 悪性ドメイン  $X_1$  から  $X_7$  として, 一般に公開されている 7 種の DGA マルウェアにより生成されたドメイン群を利用した [29, 30]. ここで, Banjori と Sisron は単語にランダムな文字列を結合することでドメインを生成するマルウェア, その他は自身が有す辞書の単語を結合することでドメインを生成するマルウェアである. それら良性と悪性ドメインに対して 5 分割交差検証を適用することで, 各データセットの 20% を検証用, 残りを学習用とした.

提案手法との比較のため, ドメインの文字列のみから悪性ドメインを判別する 2 種類の手法を実装した. それらの実装に用いた機器の構成を表 3 に示す. 第一の実装は, ドメイン文字列に対して LSTM (Long Short-Term Memory) モデルを適用することで判別する手法であり [5], 第二の実装は, マルウェアのコールバック先に基づいて構築した辞書の単語が, ドメイン文字列に閾値 2 以上含まれるか否かで判別する手法である [20]. なお, 表 4 にデータセットの 80% に基づく学習と, 残り 20% の判別に要した計算時間を示す.

提案手法における各設定は次の通りである. 単語分割機能の辞書  $\mathbb{D}$  として, Aspell [31] とコーパス [32] に登録されている単語を利用した. その単語の総数は 500,000 を超える. 単語グラフ構築機能におけるクラスタリングの閾値  $th$  を経験的に 0.85 とした. また, SVM のカーネルとして RBF (Radial Basis Function) を採用し, そのハイパーパラメータとコストを 1.0 と 5.0 とした. これらの最適化は今後の課題とする.

表 4: 各手法の計算時間

	Training	Predicting
Anderson et al. [5]	10860 s	402 s
Pereira et al. [20]	118 s	72 s
Our work	82618 s	2896 s

表 5: 実験結果

	Recall	Precision
Anderson et al. [5]	0.9977	0.9305
Pereira et al. [20]	0.8873	0.6380
Our work	0.9977	0.9869

## 4.2 定量的評価

各手法における悪意ドメインの判別性能を定量的に評価するために、一般的な2つの指標を用いた。再現率は、悪性ドメインの総数に対する悪性と判別されたドメインの数の比率であり、適合率は、悪性と判別されたドメインの総数に対する真に悪性であるドメインの数の比率である。

実験結果を表5に示す。ここで、各値は交差検証における5回の試行の平均である。この結果から、提案手法は0.9977の再現率と0.9869の適合率を達成しており、2つの実装よりも高い精度を示すことが見て取れる。2つの実装の精度が低下した理由は次の通りである。まず、文献[5]に基づく実装における、単純な文字の並びのみからドメインの良性と悪性を判別することの限界である。特筆すべきは、8文字以上のアルファベットのみから成るドメインを悪性と判別する傾向が見られ、それが故に多数の良性ドメインで誤判が生じることとなった。次いで、文献[20]に基づく実装は、ドメイン文字列における単語の重要性を画一的に測るため、良性ドメインと悪性ドメインの一部で誤判が多発したことが原因である。提案手法は、良性と悪性のドメイン文字列を成す単語の差異を考慮すること、単語グラフにおいて中心的な役割を担う単語を重視することで、これらの要因を除外できたと考えられる。

各データセットにおいて誤判したドメインの数を表6に示す。ここで、各値は交差検証における5回の試行の合計である。提案手法は、BanjoriとSisronが生成したドメインである $X_1$ と $X_6$ に対して非常に優秀な判別を実現した。その一方、良性を含む他のデータセットにおいては合計で3,241の誤判が生じる結果となった。その誤判の多くは次の4種、(a)単一の単語のみから成るドメイン、(b)学習用データセットにおいて、出現頻度が極端に少ない単語から成るドメイン、(c)一般的な単語を含むドメイン、(d)辞書に含まれない単語から成るドメインであった。本手法は、ドメイン文字列を構成する単語の関

表 6: 各データセットにおけるドメインの誤判数

	$X_0$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$
Anderson et al. [5]	15631	0	154	0	62	27	0	220
Pereira et al. [20]	105706	0	285	26	23286	28	0	36
Our work	2772	0	295	31	29	32	0	82

表 7: 提案手法と既存手法との定性的な比較

	(1) DGA Malware Detection	(2) Dict-DGA Malware Detection	(3) Robust to Encryption	(4) Network- Scale Independent	(5) Real-Time Detection
Soldo et al. [10]			✓	✓	✓
Gu et al. [12]				✓	✓
Rahbarinia et al. [15]			✓	✓	✓
Berger et al. [16]			✓		
Wang et al. [17]	✓	✓	✓		
Truong et al. [4]	✓		✓	✓	✓
Anderson et al. [5]	✓		✓	✓	✓
Pereira et al. [20]		✓	✓	✓	✓
Our work		✓	✓	✓	

係性に着目しており、良性と悪性の判別には教師有り機械学習アルゴリズムを利用している。それらの特性上、(a) と (b) のドメインの正確な判別は困難となる。(c) のドメイン文字列を調査したところ、domain, network, host, local など、良性と悪性を問わず頻出する単語を含んでいた。そのため、自然言語処理におけるストップワードを参考に [33]、ドメインにおける一般的な単語を除外することで改善が期待できる。(d) の誤判は、辞書における語彙数の不足により、ドメインを無意味な短い文字列に分割することが原因であった。具体的には、略語や頭字語を含むドメイン、非アルファベットをアルファベット表記したドメイン、固有名詞から成るドメインなどの単語分割で顕著な誤りが見られた。その改善には、より網羅的な辞書の準備が必要となる。

以上の議論より、幾つかの課題があるにしても提案手法が 0.9977 の再現率と 0.9869 の適合率で悪性ドメインを判別できることを確認した。この結果は、これら悪性ドメインの名前解決を予兆として、ネットワークに内在する辞書に基づく DGA マルウェアを高精度で検出できることを示唆している。

### 4.3 定性的評価

表 7 に提案手法と既存手法との定性的な比較を示す。その比較の観点は、(1) DGA マルウェアの検出性能、(2) 辞書に基づく DGA マルウェアの検出性能、(3) 暗号化に対する頑健性、(4) ネットワークの規模に対する依存の有無、(5) 検出の実時間性であり、各項目における対応の可否をレ点の有無により記している。なお、表中の評価は、主に 2.2 節における議論を取り纏めたものとなっている。

前節で述べたように、提案手法は辞書に基づく DGA マルウェアの高精度な検出を実現している。その検出の特徴として、Gu らの BotHunter [12] とは異なり、通信の暗号化による制限を受けないこと、Berger らの DNSMap [16] や Wang らの DBod [17] とは異なり、大規模なネットワークの観測を必要としないことが挙げられる。その一方で、一般的な DGA マルウェアのための検出機能を有しない。これは辞書に基づく DGA マルウェアに特化しているためであり、それ故に実際の運用では Anderson らの手法 [5] などを用いた補完が必須となる。

他のドメイン文字列に基づく検出と比べ [4, 5, 20]、提案手法は計算時間を要する仕組みとなっている。これは 3.3 節で述べたドメイン文字列の単語分割において、最適な候補を総当りで見つけ出すことが原因である。この緩和のため、DNS の名前解決要求がドメイン名からアドレスへの変換であること、その結果として NXDOMAIN を応答することの、両条件を満たすもののみを対象とすることでドメイン

の判別数を大きく絞り込んでいる。ここで、NXDOMAIN 応答はドメインの未登録を意味するため、その名前解決に起因したデータ通信は発生しないことを留意されたい。故に、本手法による判別には厳格な実時間性は要求されず、その実装の最適化を図るのみで運用に耐え得る性能を達成可能であると考えられる。

以上の議論により、辞書に基づく DGA マルウェアの検出における提案手法の優位性を明らかにした。具体的には、通信の暗号化やネットワークの規模に制限されることなく、感染端末の迅速な排除が可能であり、それ故にネットワークの運用における安全性への貢献が期待できる。

## 5 おわりに

本稿では、DNS に対する膨大な数の名前解決要求から、辞書に基づく DGA マルウェアにより生成されたドメインの判別を試みた。その実現に向け、ドメインの文字列を構成する単語の関係性に基づく悪性ドメイン判別手法を提案した。また実験を通じて、提案手法が 0.9977 の再現率と 0.9869 の適合率で悪性ドメインを判別可能であること、それ故に、悪性ドメインの名前解決を予兆として辞書に基づく DGA マルウェアを高精度で検出できることを確認した。この結果から、ネットワークに内在する多様なマルウェアへの迅速な対処が可能となるため、ネットワークの運用において安全性の向上が期待できる。今後は、大規模なネットワークで観測した通信を対象に、判別精度の継続的な評価を予定している。

**謝辞** 本研究は JSPS 科研費 JP21K11848 の助成を受けたものである。ここに深く謝意を示す。

## 参考文献

- [1] J. A. Lewis: Economic Impact of Cybercrime — No Slowing Down, <https://www.csis.org/analysis/economic-impact-cybercrime> (accessed 2020-09-10).
- [2] Y. Fu et al.: Stealthy Domain Generation Algorithms, *IEEE Transactions on Information Forensics and Security*, Vol. 12, No. 6, pp. 1430–1443 (2017).
- [3] A. Satoh et al.: Estimating the Randomness of Domain Names for DGA Bot Callbacks, *IEEE Communications Letters*, Vol. 22, No. 7, pp. 1378–1381 (2018).
- [4] D. Truong et al.: Detecting Domain-Flux Botnet based on DNS Traffic Features in Managed Network, *Security and Communication Networks*, Vol. 9, No. 14, pp. 2338–2347 (2016).
- [5] H. S. Anderson et al.: DeepDGA: Adversarially-Tuned Domain Generation and Detection, *Proceedings of the ACM Workshop on Artificial Intelligence and Security*, pp. 13–21 (2016).
- [6] A. K. Sood et al.: A Taxonomy of Domain-Generation Algorithms, *IEEE Security & Privacy*, Vol. 14, No. 4, pp. 46–53 (2016).
- [7] D. Kim: Potential Risk Analysis Method for Malware Distribution Networks, *IEEE Access*, Vol. 7, pp. 185157–185167 (2019).
- [8] C. Dwyer et al.: Malvertising — A Rising Threat to the Online Ecosystem, *Journal of Information Systems Applied Research*, Vol. 10, No. 3, pp. 29–37 (2017).

- [9] D. Andriess et al.: Highly Resilient Peer-to-Peer Botnets Are Here: An Analysis of GameOver Zeus, *Proceedings of the International Conference on Malicious and Unwanted Software*, pp. 116–123 (2013).
- [10] F. Soldo et al.: Blacklisting Recommendation System: Using Spatio-Temporal Patterns to Predict Future Attacks, *IEEE Journal on Selected Areas in Communications*, Vol. 29, No. 7, pp. 1423–1437 (2011).
- [11] J. Freudiger et al.: Controlled Data Sharing for Collaborative Predictive Blacklisting, *Proceedings of International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, pp. 327–349 (2015).
- [12] G. Gu et al.: BotHunter: Detecting Malware Infection Through IDS-Driven Dialog Correlation, *Proceedings of the USENIX Conference on Security Symposium*, pp. 167–182 (2007).
- [13] T. J. Parvat et al.: Performance Improvement of Deep Packet Inspection for Intrusion Detection, *Proceedings of the IEEE Global Conference on Wireless Computing & Networking*, pp. 224–228 (2014).
- [14] Cisco Systems, Inc.: Cisco 2018 Annual Cybersecurity Report, <https://www.cisco.com/c/en/us/products/security/cybersecurity-reports.html> (accessed 2020-09-10).
- [15] B. Rahbarinia et al.: Efficient and Accurate Behavior-Based Tracking of Malware-Control Domains in Large ISP Networks, *ACM Transactions on Privacy and Security*, Vol. 19, No. 2, pp. 4:1–4:31 (2016).
- [16] A. Berger et al.: Mining Agile DNS Traffic Using Graph Analysis for Cybercrime Detection, *Computer Networks*, Vol. 100, pp. 28–44 (2016).
- [17] T. S. Wang et al.: DBod: Clustering and Detecting DGA-based Botnets using DNS Traffic Analysis, *Computers & Security*, Vol. 64, pp. 1–15 (2017).
- [18] D. Plohmann et al.: A Comprehensive Measurement Study of Domain Generating Malware, *Proceedings of the USENIX Conference on Security Symposium*, pp. 263–278 (2016).
- [19] R. Vinayakumar et al.: Evaluating Deep Learning Approaches to Characterize and Classify the DGAs at Scale, *Journal of Intelligent and Fuzzy Systems*, Vol. 34, No. 3, pp. 1265–1276 (2018).
- [20] M. Pereira et al.: Dictionary Extraction and Detection of Algorithmically Generated Domain Names in Passive DNS Traffic, *Proceedings of the International Symposium on Research in Attacks, Intrusions, and Defenses*, pp. 295–314 (2018).
- [21] A. Koren: Ursnif Malware: Deep Technical Dive, <https://arielkoren.com/blog/2016/11/01/ursnif-malware-deep-technical-dive/> (accessed 2020-09-10).
- [22] S. Skuratovich: Matsnu: A Deep Dive, <https://blog.checkpoint.com/2015/07/02/matsnu-a-new-malware-discovery/> (accessed 2020-09-10).
- [23] D. Sahoo et al.: Malicious URL Detection using Machine Learning: A Survey, *arXiv:1701.07179*, pp. 1–21 (2017).
- [24] P. Mockapetris: Domain Names — Implementation and Specification, IETF Request for Comments 1035 (1987).

- [25] A. Costello: Punycode: A Bootstring Encoding of Unicode for Internationalized Domain Names in Applications (IDNA), IETF Request for Comments 3492 (2003).
- [26] D. Müllner: fastcluster: Fast Hierarchical, Agglomerative Clustering Routines for R and Python, *Journal of Statistical Software*, Vol. 53, No. 9, pp. 1–18 (2013).
- [27] G. Csárdi et al.: The igraph Software Package for Complex Network Research, *InterJournal Complex Systems*, No. 1695 (2006).
- [28] A. Karatzoglou et al.: Support Vector Machines in R, *Journal of Statistical Software*, Vol. 15, No. 9, pp. 1–28 (2006).
- [29] J. Bader: Some Results of My DGA Reversing Efforts, [https://github.com/baderj/domain\\_generation\\_algorithms](https://github.com/baderj/domain_generation_algorithms) (accessed 2020-09-10).
- [30] Fraunhofer FKIE: DGArchive, <https://dgarchive.caad.fkie.fraunhofer.de> (accessed 2020-09-10).
- [31] K. Atkinson: GNU Aspell, <http://aspell.net> (accessed 2020-09-10).
- [32] P. Norvig: Natural Language Corpus Data: Beautiful Data, <http://norvig.com/ngrams/> (accessed 2020-09-10).
- [33] J. Nothman et al.: Stop Word Lists in Free Open-source Software Packages, *Proceedings of the Workshop for NLP Open Source Software*, pp. 7–12 (2018).



## 利用実績

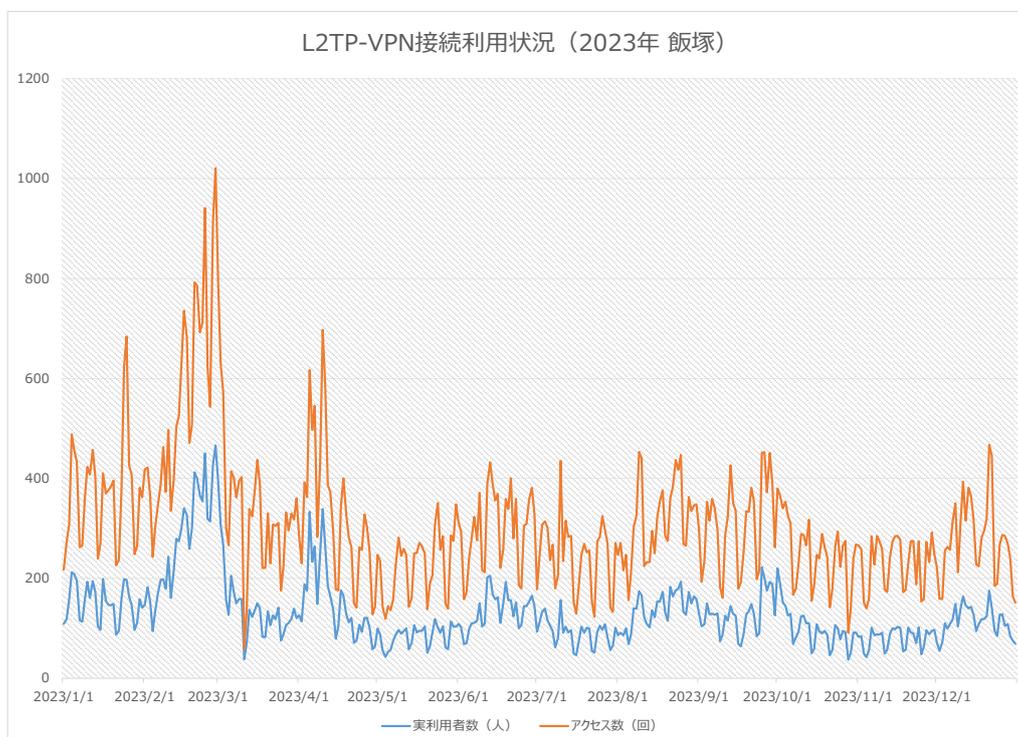
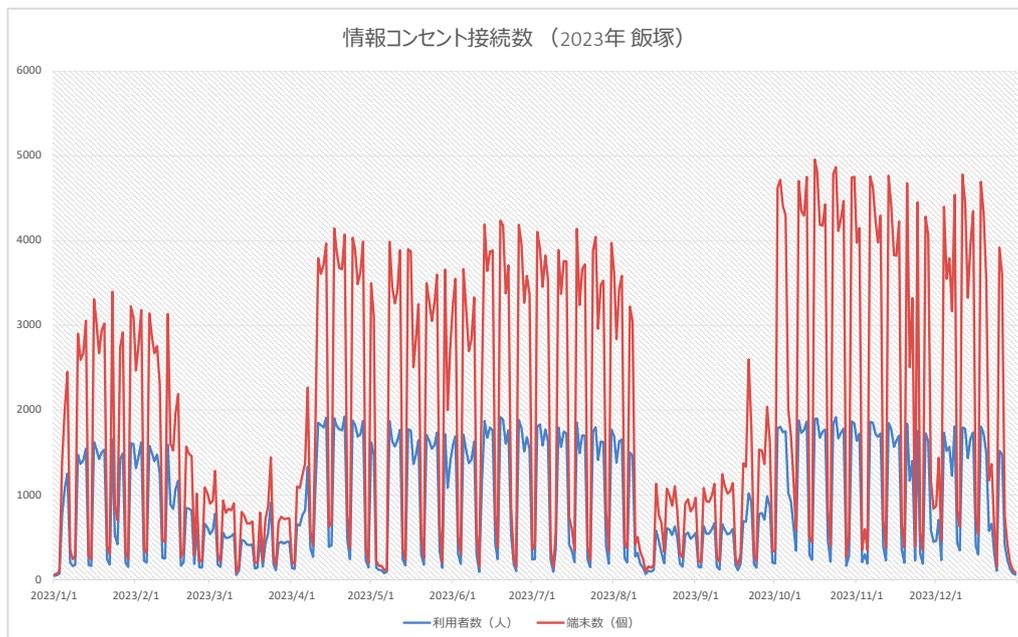
次の実績報告を示します。

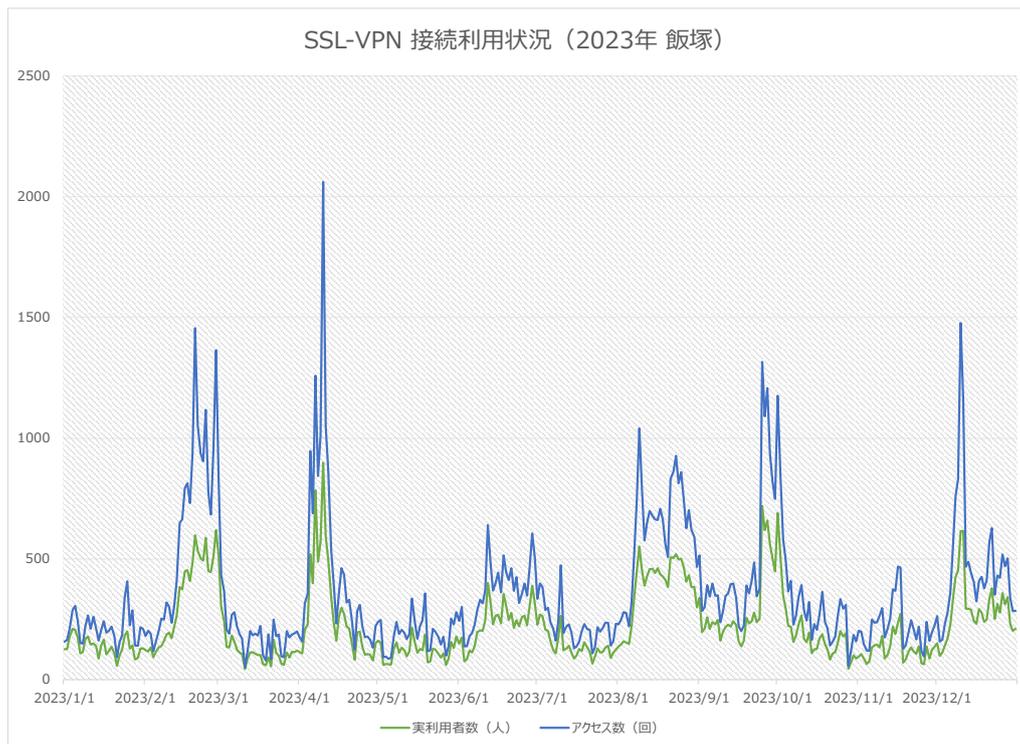
- 情報コンセント接続および VPN 接続の利用状況
- 九工大メールのアカウント発行実績
- 通常講義以外での講義室の利用状況
- 情報基盤センターへの訪問者
- 講習会の参加人数
- 各キャンパスの講義室の時間割
- 質問者の集計

# 1 情報コンセント及びVPNの利用状況

## 1.1 飯塚キャンパス

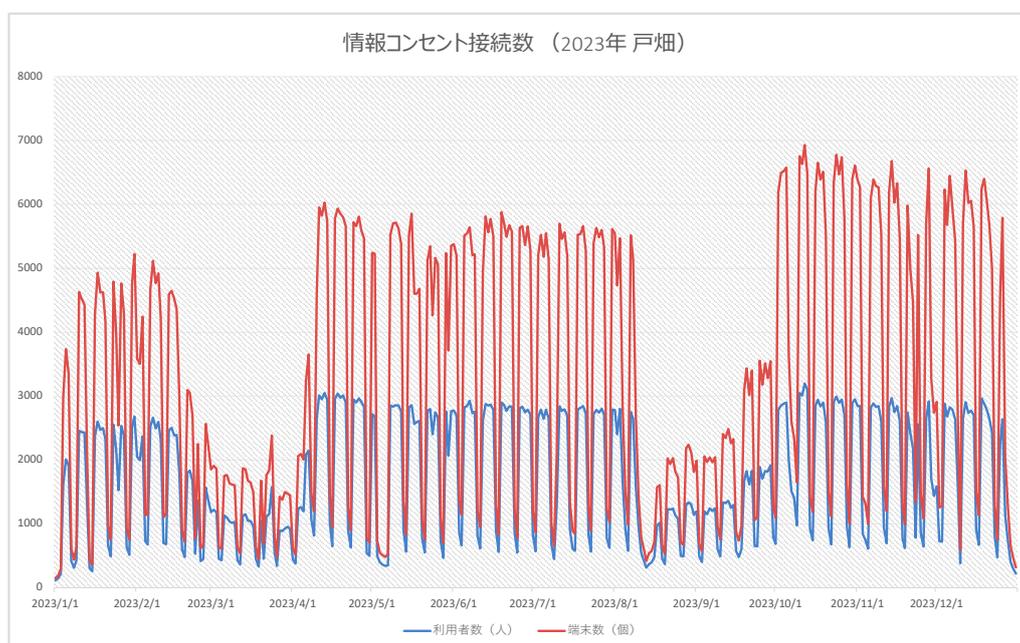
2023年1月から12月までに利用された、情報コンセント・VPNの利用状況を示します。

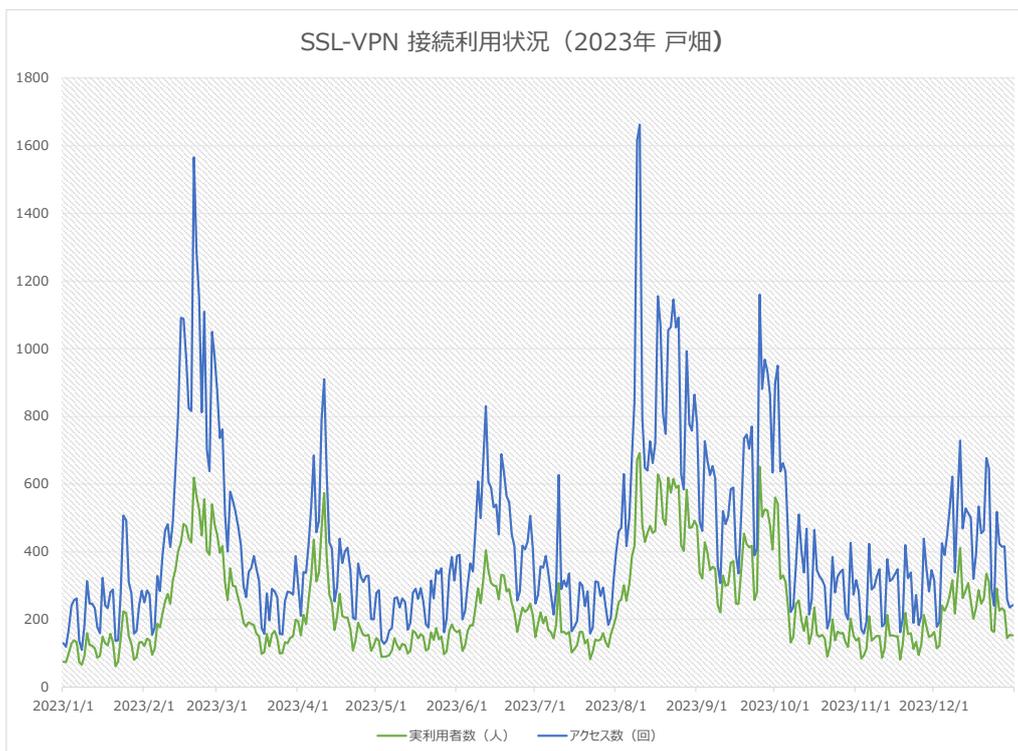
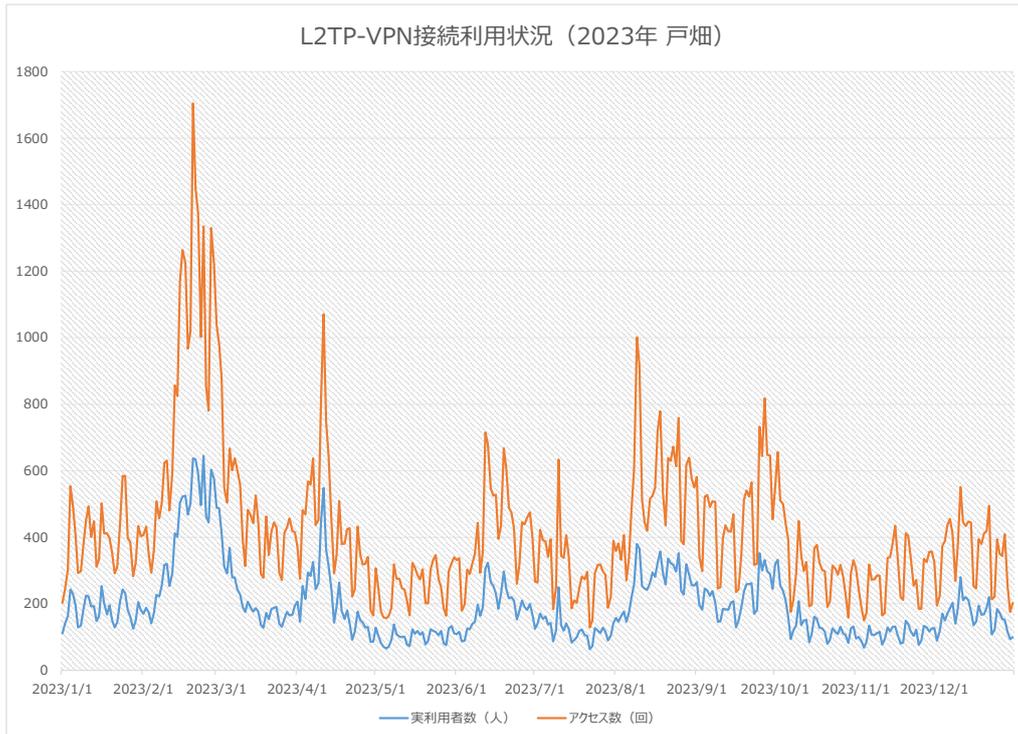




## 1.2 戸畑キャンパス

2023年1月から12月までに利用された、情報コンセント・VPNの利用状況を示します。

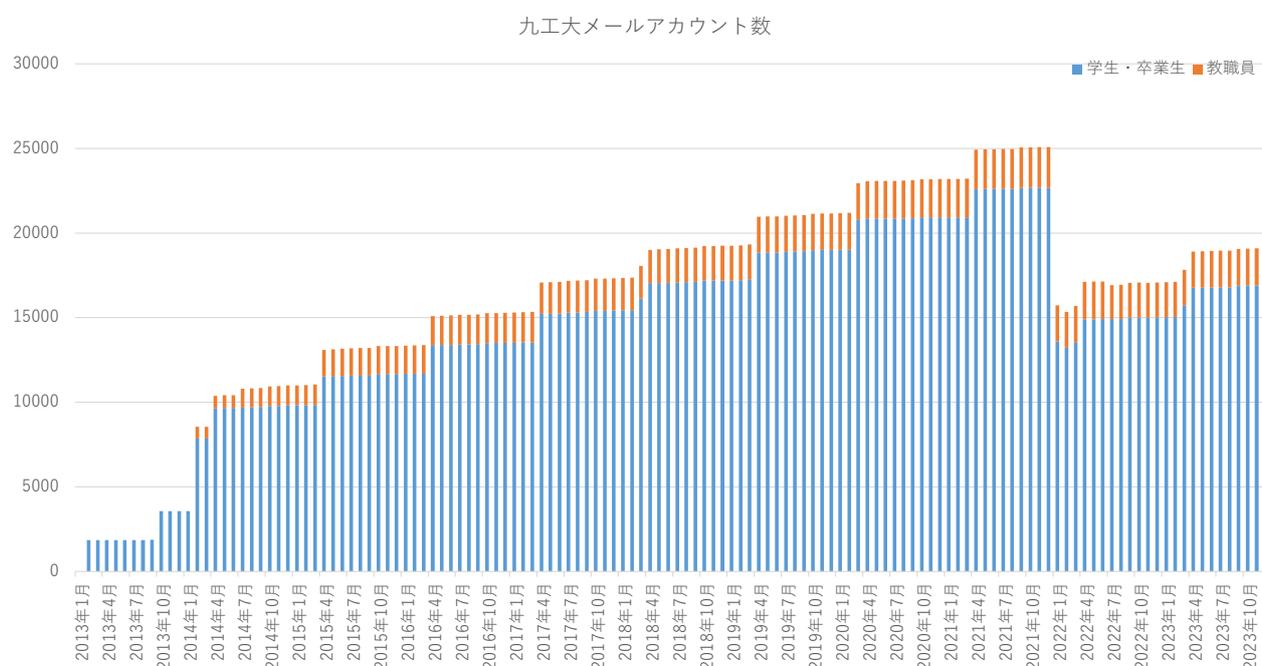




## 2 九工大メールアカウント数

### 2.1 アカウント数の推移

九工大メールサービス開始年(2013年1月)から2024年1月までの九工大メールアカウント数(学生, 教職員)の推移を示します。



2023年12月時点のアカウント数：学生(卒業生含む)16912, 教職員(退職者含む)2201

### 2.2 多数のアカウントを処理した年月一覧

多数(500アカウント以上)のアカウントを処理した年月, 数および事由は以下の通りです。

- 2013年2月** 2012年度卒業予定者向け発行(発行数:1849)
- 2013年10月** 2013年度卒業予定者向け発行(発行数:1700)
- 2014年2月** 全在学生への付与に伴う発行(発行数:4989), 全教職員への付与に伴う発行(発行数:679)
- 2014年4月** 2014年度入学生向け発行(発行数:1794)
- 2015年4月** 2015年度入学生向け発行(発行数:1696)
- 2016年4月** 2016年度入学生向け発行(発行数:1688)
- 2017年4月** 2017年度入学生向け発行(発行数:1687)
- 2018年4月** 2018年度入学生向け発行(発行数:1697)
- 2019年4月** 2019年度入学生向け発行(発行数:1585)
- 2020年4月** 2020年度入学生向け発行(発行数:1513)
- 2021年4月** 2021年度入学生向け発行(発行数:1677)
- 2022年2月** 長期間利用のない卒業生・離退職者アカウント削除(削除数:9739)
- 2022年2月** 2022年度入学生向け発行(発行数:1680)
- 2023年2月** 2023年度入学生向け発行(発行数:1717)

### 3 通常講義以外での講義室の利用状況

飯塚では2023年6月からAV講義室・AV演習室・端末講義室を情報工学部へ、戸畑では2023年3月からC-2B講義室・C-2G講義室を工学部へ移管したため飯塚は2023年5月まで、戸畑は2023年2月までの実績となります。

#### 3.1 情報基盤センター(飯塚)の講義室

内容	利用対象者	利用日時・講義室
学部オリエンテーション	学生	2023.04.06 1-5 限目 (AV 講義室)
学部オリエンテーション	学生	2023.04.07 8:50-21:10(AV 講義室)
学部オリエンテーション	学生	2023.04.07 8:50-21:10(AV 演習室)
学部オリエンテーション	学生	2023.04.07 8:50-21:10(端末講義室)

#### 3.2 情報基盤センター(戸畑)の講義室

内容	利用対象者	利用日時・講義室
講義	学生	2023.01.12 3-4 限目 (C-2B 講義室)
入試対応	職員	2023.01.12-15 終日 (C-2B 講義室)
講義	学生	2023.01.13 2 限目 (C-2G 講義室)
作業	職員	2023.01.16 3 限目 (C-2G 講義室)
作業	職員	2023.01.17 3-4 限目 (C-2G 講義室)
講義	学生	2023.01.17 6 限目 (C-2B 講義室)
作業	職員	2023.01.18 昼休み (C-2G 講義室)
講義	学生	2023.01.23 5(C-2G 講義室)
講義	学生	2023.01.24 3-5 限目 (C-2B 講義室)
講義	学生	2023.02.02 3-4 限目 (C-2G 講義室)
講義	学生	2023.02.03 1-2 限目 (C-2B 講義室)
入試対応	職員	2023.02.03-05 終日 (C-2B 講義室)
講義	学生	2023.02.08 3-4 限目 (C-2B 講義室)
講義	学生	2023.02.09 4 限目 (C-2B 講義室)
講義	学生	2023.02.16 3-5 限目 (C-2B 講義室)
作業	職員	2023.01.18-19 1-5 限目 (C-2G 講義室)
講義	学生	2023.02.21 1-6 限目 (C-2B 講義室)
講義	学生	2023.02.22 4 限目 (C-2B 講義室)
講義	学生	2023.02.22 4 限目 (C-2G 講義室)
入試対応	職員	2023.02.22-25 終日 (C-2B 講義室)
講義	学生	2023.02.24 4 限目 (C-2G 講義室)

## 4 訪問者

2023年1月から2023年12月までの間の情報基盤センターへの訪問者及び人数を、キャンパス別に示します。

### 4.1 情報基盤センター(飯塚)への訪問者

- 該当なし

### 4.2 情報基盤センター(戸畑)への訪問者

- 該当なし

## 5 講習会の参加人数

2023年1月から2023年12月までの間に開催された講習会の参加人数について、キャンパス別に示します。

### 5.1 飯塚キャンパス

- 該当なし

### 5.2 戸畑キャンパス

- 該当なし

## 6 各キャンパスの講義室の時間割

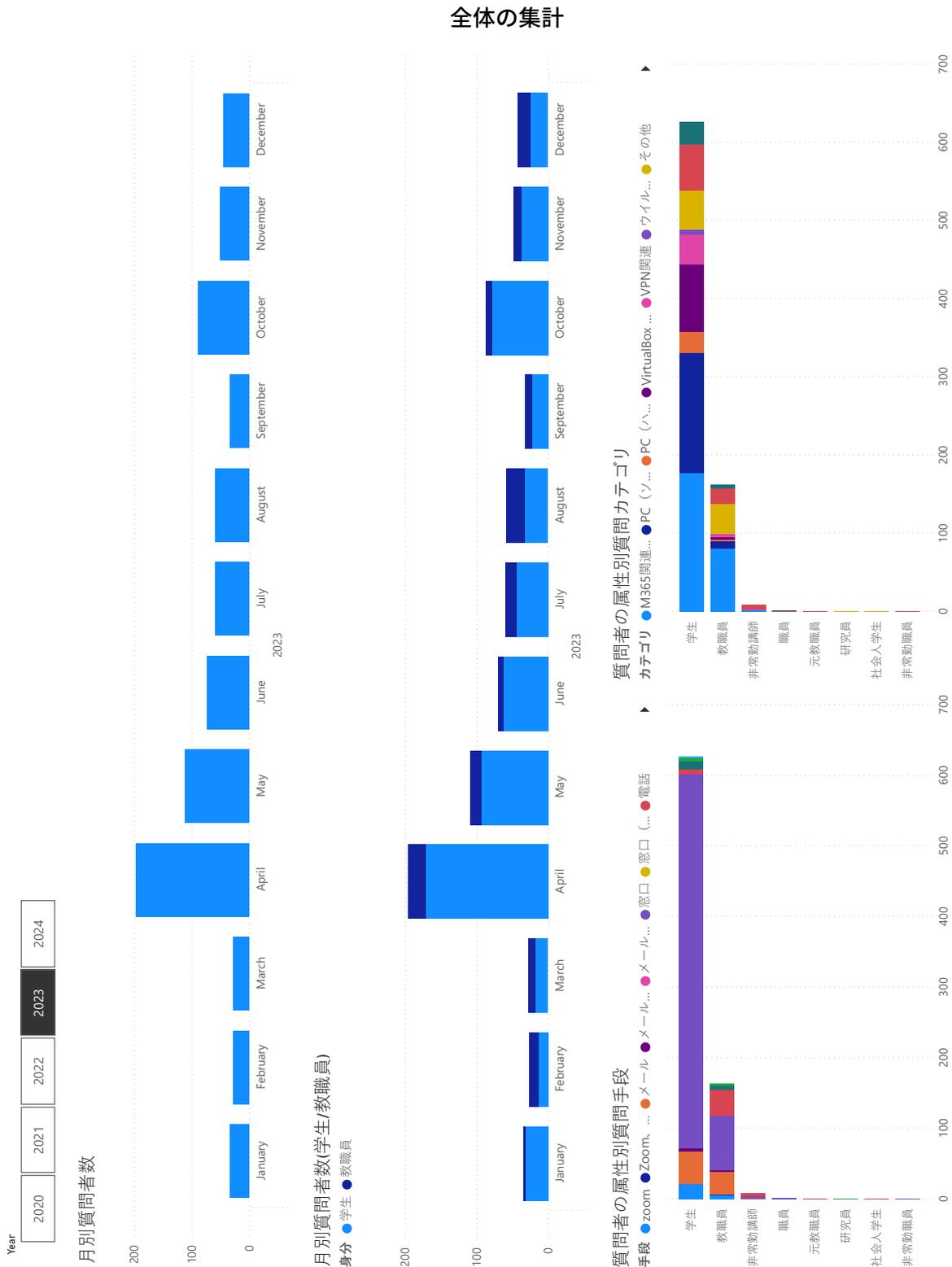
飯塚では、2023年6月からAV講義室・AV演習室・端末講義室を情報工学部へ移管したため、第1Qの時間割のみを示します。戸畑では、2023年3月からC-2B講義室・C-2G講義室を工学部へ移管したため、時間割はありません。

飯塚 2023 年度 第 1 クォーター 講義時間割

飯塚 2023 年度 第 1 クォーター 時 間 割			
		AV講義室	端末講義室
月曜日	1		
	2		
	3	プログラミング	プログラミング
	4	Ⅱ 1年	I 1年
	5	乃万	新見
火曜日	1		プログラム設計
	2		知的2年 古賀雅
	3	プログラミング	プログラム設計
	4	合同クラスB(ⅢB + ⅣB) 1年	物理2年 嶋田
	5	梅田	
水曜日	1	プログラミング	
	2	V 1年	
	3	齊藤剛	情報通信工学実験I
	4		情通2年 黒崎
	5		
木曜日	1		
	2		
	3		
	4		
	5		
金曜日	1		
	2	人工知能プログラミング	
	3	知能3年 國近	プログラム設計
	4		知能2年 碓崎
	5		

## 7 質問者の集計

2023年1月から2023年12月までの間に、情報基盤センター窓口に来た質問者の集計を示します。

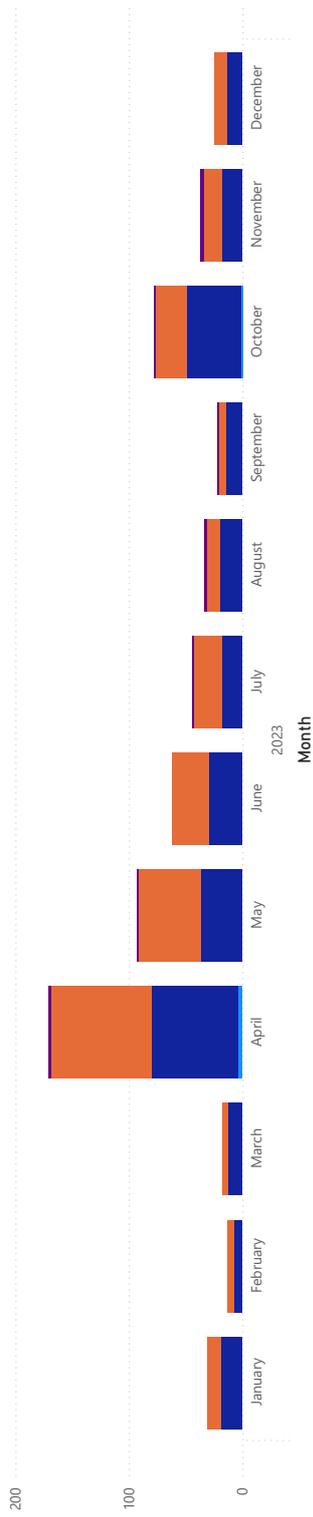






月別質問者数

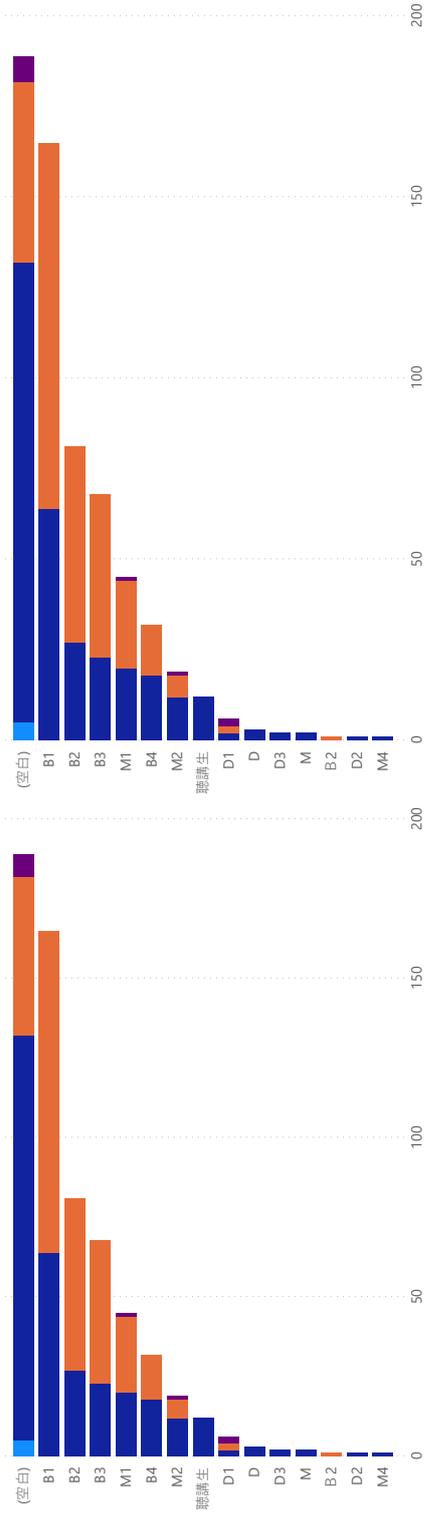
学科 ● (空白) ● 工学部 ● 情報工学部 ● 生命体



学生のみの集計

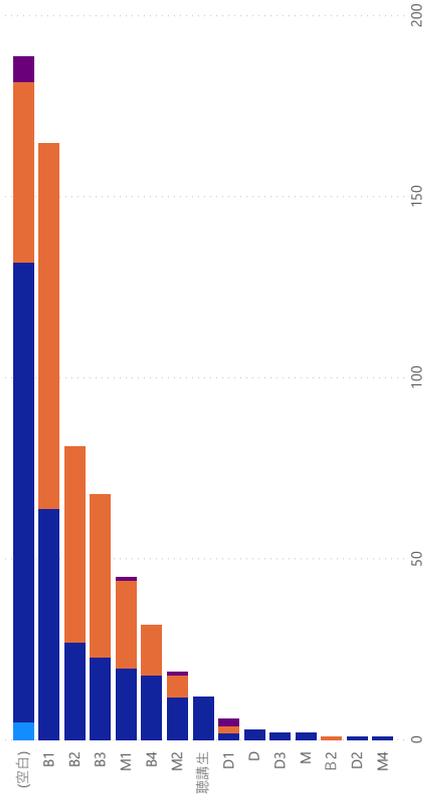
学年別質問者数

学科 ● (空白) ● 工学部 ● 情報工学部 ● 生命体



キャンパス別質問者数

学科 ● (空白) ● 工学部 ● 情報工学部 ● 生命体



カテゴリ別の集計

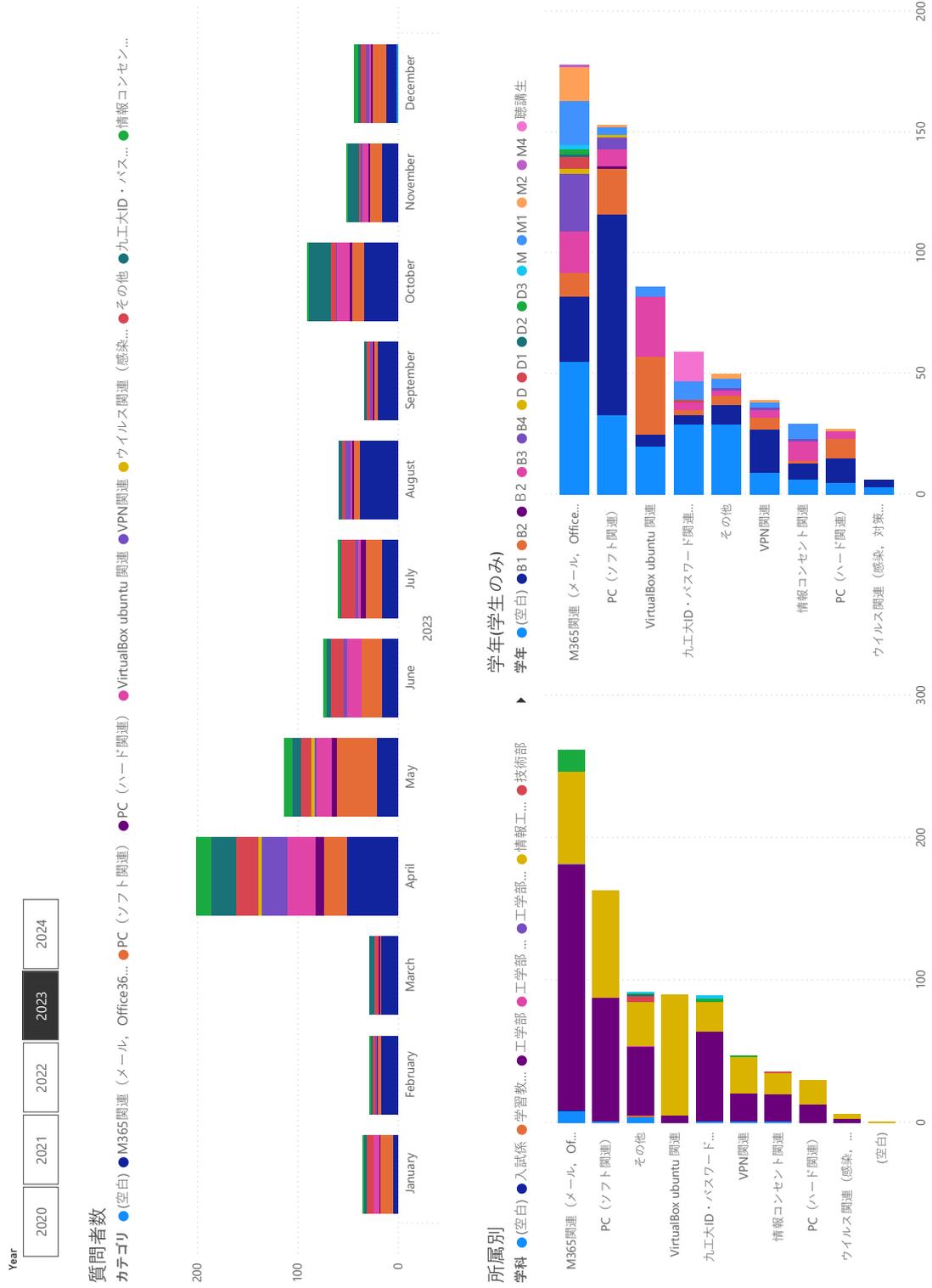




表 2: 支援を行った学会・研究会等一覧 (情報コンセント利用)

期日	キャンパス	行事名	支援内容
令和5年5月13日	戸畑	九州工業大学技術士会	発行数:20
令和5年8月19日 ～ 令和5年8月20日	飯塚	免許法認定通信教育 教科教育法 (情報) II	発行数:10
令和5年8月26日	飯塚	小学生対象プログラミング教室	発行数:15
令和5年9月6日 ～ 令和5年9月8日	戸畑	DV-X α 研究会	発行数:50
令和5年10月15日	飯塚	サイレント音声認識ワークショップ	発行数:30
令和5年10月21日	飯塚	ISG フェスタ	発行数:10
令和5年10月21日 ～ 令和5年10月22日	戸畑	日本細胞性粘菌学会例会	発行数:30
令和5年10月26日	飯塚	大学見学	発行数:25
令和5年12月11日 ～ 令和5年12月17日	飯塚	さくらサイエンスプラン	発行数:10
令和5年12月14日 ～ 令和5年12月15日	飯塚	電子情報通信学会 ネットワークシステム研究会	発行数:50
令和6年2月20日	飯塚	学術・教育コンテンツ共有流通部会 (AXIES-CSD) 研究会	発行数:20







## 報告 (本年度の活動)

- 佐藤彰洋, 戸田哲也, 和田数字郎, 福田豊, 中村豊, “学外公開アドレス管理システム”, Vol.27, No.1, pp.167-173, Nov 2023.
- 林豊洋, 黒崎覚, 金光昂志, “九州工業大学における Microsoft Teams の全学展開”, 大学 ICT 推進協議会 2023 年度年次大会論文集, pp.97-102, Dec 2023.
- 山口真之介, 大西淑雅, 西野和典, “非同期型情報リテラシー講義における学習活動と効果の分析”, 大学 ICT 推進協議会 2023 年度年次大会論文集, pp.297-303, Dec 2023.
- 大西淑雅, “API 連携の実例紹介”, Moodle、Mahara プラグイン、LTI Tool 作成ワークショップ, Feb 2024.
- 大西淑雅, “学習活動データを学習者に提供するプラグインの提案”, Moot Moot Japan 2024, Feb 2024.
- 尋木信一, 山田雅之, 鈴木計哉, 大西淑雅, 山口真之介, 浅羽修丈, 西野和典, “高校情報科教員を対象とする授業設計支援システムー情報科教員養成課程および現職教員による試用評価ー”, 日本情報科教育学会第 22 回研究会論文集, pp.27-32, Mar 2024.
- 林豊洋, 黒崎覚, “部局内メールアドレスの Microsoft365 への集約について”, 九州工業大学情報基盤センター年報, No.4, pp.3-10, Mar 2024.
- 大西淑雅, 山口真之介, “ビデオ会議サービスの提供と API 連携機能の導入”, 九州工業大学情報基盤センター年報, No.4, pp.11-24, Mar 2024.

## 4 研究関連 (外部資金獲得等)

- 科学研究費 基盤研究 (C) (研究代表者) (中村)
- 科学研究費 基盤研究 (C) (研究代表者) (大西)
- 科学研究費 基盤研究 (C) (研究代表者) (佐藤)

## 5 研究関連 (受賞等)

- 三島和宏, 中村豊, 福田豊, 柏崎礼生, 中村素典, 森村吉貴, 北口善明: 情報処理学会インターネットと運用技術研究会 藤村記念ベストプラクティス賞

## 6 社会貢献関連 (講演等)

- 大西淑雅: 大学 ICT 推進協議会 (AXIES)・年次大会 2023 企画セッション「15AM2C Learning-Analytics におけるオープンソースソフトウェアの活用」オーガナイザ
- 大西淑雅: 大学 ICT 推進協議会 (AXIES)/大学 e ラーニング協議会・JSISE 全国大会プレカンファレンス「Moodle の機能拡張と活用実践」企画・実施
- 大西淑雅: 大学 ICT 推進協議会 (AXIES)・「OSS Cafe 2023 ～学習支援システム (LMS: Learning Management System) の運用～」企画・発表

- 大西淑雅：大学eラーニング協議会・UeLA フォーラム「理工系総合大学におけるLMS運用の現状と展望」企画・実施・オーガナイザ

## 7 社会貢献関連(理事, 委員等)

- 九州大学情報基盤研究開発センター計算委員会委員, 情報処理学会インターネットと運用技術研究会運営委員, 情報処理学会第16回インターネットと運用技術シンポジウムプログラム委員(林)
- 情報処理学会インターネットと運用技術研究会運営委員, 情報処理学会第16回インターネットと運用技術シンポジウムプログラム委員・実行委員, 情報処理学会論文誌特集号編集委員(福田)
- ロボカップ日本委員会理事, ロボカップジュニア・ジャパン代表理事, ロボカップフェデレーションジャパン監事(大橋)
- 情報処理学会・査読委員, 大学ICT推進協議会(AXIES)オープンソース技術部会運営委員, 大学eラーニング協議会第一部会部会長, 大学ICT推進協議会(AXIES)年次大会2023論文審査委員・プログラム委員, 大学学習資源コンソーシアム会員, 日本IMS協会地域会員, 日本オープンオンライン教育推進協議会(JMOOC)個人会員(大西)
- 情報処理学会モバイルコンピューティングと新社会システム研究会特任委員, 電子情報通信学会インターネットアーキテクチャ研究会専門委員, 情報処理学会論文誌「移動の価値を再創造する高度交通システムとパーベイシブシステム」特集号編集委員(佐藤)

## 8 学内委員, 情報システム仕様策定・技術審査

- 「全学ICT教育研究基盤システム」仕様策定委員, 「教育研究用計算機基盤システム(生命体工学研究科)」仕様策定委員, 「附属図書館学術情報システム」仕様策定委員, 「情報工学教育研究用コンピュータシステムライセンス更新」仕様策定委員, 「教務システムの稼働を目的とした事務用仮想基盤システム増設」技術審査員(林)
- 「全学ICT教育研究基盤システム」仕様策定委員, 学習教育センター・教育DX支援グループ・兼任(大西)

## センター人事異動および職員配置

### 1 人事異動

2023年1月から2023年12月までのセンター人事異動を示す。

令和5年	4月	1日	教授	鶴 正人	センター長退任
令和5年	4月	1日	教授	中村 豊	センター長就任

### 2 センター職員配置

2023年1月現在のセンター職員の配置を示す。なお、その他にも学部生よりなる技術補佐員が配置されている。

	職名	氏名	主な勤務地	連絡先1	連絡先2
センター長	教授	中村 豊	戸畑	戸畑 3472	飯塚 7555
副センター長	准教授	林 豊洋	飯塚	飯塚 7551	————
	教授	大橋 健	飯塚	飯塚 7569	————
	准教授	大西 淑雅	飯塚	飯塚 7571	————
	〃	福田 豊	戸畑	戸畑 3474	————
	〃	佐藤 彰洋	戸畑	戸畑 3473	————
	助教	中山 仁	飯塚	飯塚 7552	————
	技術専門職員	井上 純一	飯塚	飯塚 7558	————
	〃	富重 秀樹	飯塚	飯塚 7558	————
	〃	戸田 哲也	戸畑	戸畑 3471	————
	〃	和田 数字郎	戸畑	戸畑 3471	————
	〃	畑瀬 卓司	戸畑	戸畑 3471	————
	事務補佐員	青木 文子	戸畑	戸畑 3470	戸畑 3471
	〃	坂口 久美	飯塚	飯塚 7555	————
	技術補佐員	辻田 尚子	飯塚	飯塚 7558	飯塚 7555
	〃	杉町 妙子	戸畑	戸畑 3470	戸畑 3471

◇◇◇◇◇  
利用規則  
◇◇◇◇◇

## 情報科学センター規則等

情報科学センターに関連する以下の規則等，加えて九州工業大学情報システム利用規程を示す。

- 九州工業大学情報基盤センター規程
- 九州工業大学情報基盤センター利用細則
- 九州工業大学ネットワークセキュリティ基盤運用室規程
- 九州工業大学 ICT 教育研究基盤運用室規程

# 九州工業大学情報基盤センター規程

令和 2年 3月 9日

九工大規程第 3号

改正 令和4年7月27日九工大規程第18号

## ○九州工業大学情報基盤センター規程

(目的)

**第1条** この規則は、九州工業大学情報基盤機構規則（平成25年九工大規則第1号）第3条の規定に基づき、九州工業大学情報基盤センター（以下「センター」という。）に関し、必要な事項を定めることを目的とする。

(業務)

**第2条** センターは、次の業務を行う。

- (1)DX 推進室の業務
- (2)ネットワークセキュリティ基盤運用室の業務
- (3)ICT 利活用教育研究基盤運用室の業務
- (4)情報科学に関する研究開発
- (5)その他センターに関し必要な業務

(組織)

**第3条** センターに、次に掲げる職員を置く。

- (1)センター長
- (2)副センター長
- (3)その他必要な職員

(センター長)

**第4条** センター長は、情報統括副本部長をもって充てる。

**第5条** 副センター長は、センター専任の教授又は准教授の中から情報基盤機構長が任命する。

2 副センター長は、センター専任の教授又は准教授の中から情報基盤機構長が任命する。

(管理運営等の審議)

**第6条** センターの管理運営等に関する審議は、九州工業大学情報統括本部運営会議において行う。

(雑則)

**第7条** の規則に定めるもののほか、必要な事項は、別に定める。

附 則

- 1 この規則は、令和 2年4月1日から施行する。
- 2 九州工業大学情報科学センター規則（平成26年3月5日九工大規則第5号）は廃止する。
- 3 九州工業大学情報科学センターに関する専門委員会要項（平成19年情報科学センター長裁定）は、廃止する。

附 則

この規程は、令和4年7月27日から施行し、令和4年4月1日から適用する。

# 九州工業大学情報基盤センター利用細則

令和 2年 3月 9日  
九工大細則第 5号

## ○九州工業大学情報基盤センター利用細則

(目的)

**第1条** この細則は、九州工業大学情報基盤センター規程（令和2年九工大規則第 号）第8条の規定に基づき、九州工業大学情報基盤センター（以下「センター」という。）の利用に関し、必要な事項を定めることを目的とする。

(利用の原則)

**第2条** センターの利用は、教育、研究、教育研究支援その他九州工業大学（以下「本学」という。）の運営上必要と認められるものに限るものとする。

(利用の資格)

**第3条** センターを利用することができる者は、次のとおりとする。

- (1) 本学に所属する職員及び学生
- (2) 情報基盤センター長（以下「センター長」という。）が特に許可した者

(利用の承認)

**第4条** センターを利用しようとする者は、センター長の承認を受けなければならない。

(目的外利用の禁止)

**第5条** センターの利用の承認を受けた者は、承認を受けた利用目的以外に利用し、又は他人に使用させてはならない。

(利用状況の届出等)

**第6条** 利用者は、センターの利用を終了し、又は中止したときは、速やかにセンター長に届け出なければならない。

(損害賠償)

**第7条** 利用者が、故意又は重大な過失により設備等を損傷したときは、その損害に相当する費用を負担しなければならない。

(利用の取消)

**第8条** センター長は、利用者がこの細則に違反し、又はセンターの運営に重大な支障を生じ

させたときは、その利用の承認を取消し、又はその利用を停止することができる。

(経費の負担)

**第9条** センターの利用にあたっては、利用に係る経費の一部を負担しなければならない。ただし、センター長が特に必要があると認めたときは、利用経費の一部又は全部を免除することができる。

(情報システム利用規程の遵守)

**第10条** 利用者は九州工業大学情報システム利用規程（平成20年九工大規程第22号）を遵守しなければならない。

(雑則)

**第11条** この規程に定めるもののほか、センターの利用に関し必要な事項は、別に定める。

附 則

- 1 この規程は、令和2年4月1日から施行する。
- 2 九州工業大学情報科学センター利用規程(昭和63年九工大規程第21号)は、廃止する。

# 九州工業大学ネットワークセキュリティ基盤運用室規程

令和 2年 3月 9日

九工大規程第 3号

改正 令和4年7月27日九工大規程第18号

## ○九州工業大学ネットワークセキュリティ基盤運用室規程

(趣旨)

**第1条** この規程は、国立大学法人九州工業大学情報統括本部規程（令和4年九工大規程第18号）第14条の規定に基づき、ネットワークセキュリティ基盤運用室（以下「運用室」という。）の業務及び構成等に関し必要な事項について定めるものとする。

(業務)

**第2条** 運用室は、次に掲げる業務を行う。

(1) 学外ネットワークへの接続及び学内情報ネットワーク並びにそれらを構成する機器等の運用管理に関すること。

(2) 学内情報ネットワークに係る資源割当及びサブネットワークの申請等に関すること。

(3) 学内サブネットワークの技術支援に関すること。

(4) 情報セキュリティの確保及び情報セキュリティ・インシデント対応に関すること。

(5) 情報セキュリティ・インシデントの発生時に初動対応として行う学内情報ネットワーク接続からの強制的な遮断に関すること。

(6) 情報機器のデジタル・フォレンジック（物理的なアクセス、持ち帰り、証拠保全、調査及び個人情報を含むログの解析等）の運用管理に関すること。

(7) 前各号に係る学内組織との連絡及び協力並びに支援等の調整に関すること。

(8) その他全学ネットワーク基盤および情報セキュリティ対策の運用に関すること。

(構成)

**第3条** 運用室に、次の室員を置く。

(1) 情報統括本部長が指名する者

(2) 情報基盤センターの教育職員若干名

(3) 工学研究院情報基盤室長

(4) 情報工学研究院情報基盤室の室長又は副室長1名

(5) 生命体工学研究科の情報通信基盤担当教育職員1名

(6) 工学研究院情報基盤室及び情報工学研究院情報基盤室の技術職員各1名

(7) 技術部技術二課システム開発系の技術職員1名

(8) 情報基盤課の事務職員

(9) その他情報統括本部長が推薦する者若干名

(室長)

**第4条** 室長は、第3条の室員の中から、情報統括本部長が指名する者をもって充て、運用室の業務を総括する。

2 室長に事故があるときは、あらかじめ室長の指名する者が室長の職務を代行する。

(副室長)

**第5条** 副室長は、情報統括本部長が指名する者をもって充て、室長を補佐する。

(部会)

**第6条** 室長は、必要に応じて運用室に部会を置くことができる。

附 則

1 この規程は、令和2年4月1日から施行する。

2 九州工業大学情報基盤運用室規則（平成25年3月6日九工大規則第1号）は廃止する。

附 則

この規程は、令和4年7月27日から施行し、令和4年4月1日から適用する。

# 九州工業大学 ICT 利活用教育研究基盤運用室規程

令和 2年 3月 9日

九工大規程第 3号

改正 令和4年7月27日九工大規程第18号

## ○九州工業大学 ICT 利活用教育研究基盤運用室規程

(趣旨)

**第1条** この規程は、国立大学法人九州工業大学情報統括本部規程（令和4年九工大規程第18号）第14条の規定に基づき、ICT利活用教育研究基盤運用室（以下「運用室」という。）の業務及び構成等に関し必要な事項について定めるものとする。

(業務)

**第2条** 運用室は、次に掲げる業務を行う。

- (1) 必携ノートパソコンに関する学習・教育環境の構築・運用管理に関すること。
- (2) 情報関連及び情報利活用教育の支援に関すること。
- (3) 研究支援サービスの提供支援に関すること。
- (4) 全学統合ID管理システムの構築及び技術的運用に関すること。ただし、全学認証を必要とする学内サービスに関する調整も含む。
- (5) 九工大メールサービスの構築および技術的運用に関すること。
- (6) ICT利活用教育研究基盤に係る文書作成、広報に関すること。
- (7) 前各号に係る学内組織との連絡および協力並びに支援等の調整に関すること。
- (8) その他 ICT利活用教育研究基盤の運用に関すること。

(構成)

**第3条** 運用室に、次の室員を置く。

- (1) 情報統括本部長が指名する者
- (2) 情報基盤センターの教育職員若干名
- (3) 情報基盤課の事務職員
- (4) その他情報統括本部長が推薦する者若干名

(室長)

**第4条** 室長は、第3条の室員の中から、情報統括本部長が指名する者をもって充て、運用室の業務を総括する。

2 室長に事故があるときは、あらかじめ室長の指名する者が室長の職務を代行する。

(部会)

**第5条** 室長は、必要に応じて運用室に部会を置くことができる。

附 則

この規程は、令和2年4月2日から施行する。

附 則

この規程は、令和4年7月27日から施行し、令和4年4月1日から適用する。