情報処理学会論文誌 Vol.58 No.3 695–707 (Mar. 2017)

SSHパスワードクラッキング攻撃における データサイズを用いる検知手法の提案と運用評価

清水 光司[1,a)] 小刀稱 知哉[1,†1] 池部 実[2] 吉田 和幸[3]

受付日 2016年6月27日, 採録日 2016年12月1日

概要：インターネットを利用した不正アクセスが多く発生している．特に，SSH サーバに対する不正アクセスの件数は依然として多い．そこで，我々は SSH へのパスワードクラッキング攻撃を検知することを目的として「SSH パスワードクラッキング攻撃検知システム（SCRAD)」を開発・運用してきた．本システムでは SSH サーバと送信元間の 1 コネクションのパケット送受信回数からパスワードクラッキング攻撃を検知している．従来のシステムの運用結果を分析したところ，scp や rsync により少量のデータを送受信する際に正規ユーザを誤検知していた．scp によるコネクションを調査して，パケット数は攻撃者コネクションと類似しているが，データサイズは攻撃者コネクションよりも大きい傾向にあると判明した．本論文では，パケット数による検知手法の誤検知を改善するために，データサイズを用いる検知手法を提案する．また，2014 年 7 月に収集したパケットデータを入力として提案手法と従来のパケット数を用いる検知手法を比較し，提案手法の有用性を検証した結果，攻撃者のコネクション判定率を維持しつつ正規ユーザのコネクション判定率を 67.3%から 90.1%に改善できた．さらに，2015 年 2 月から 2016 年 5 月までの提案手法の運用結果を分析した結果，攻撃者のコネクションを 99.7%正確に判定できたが，正規ユーザのコネクション判定率は 72.2%であった．提案手法により SSH パスワードクラッキング攻撃による SSH サーバへの侵入のリスクを低減できることと，正規ユーザを誤検知することにより，ユーザの利便性を損ねる可能性があることが判明した．

キーワード：SSH，パスワードクラッキング攻撃，TCP，アノマリ型不正通信検知

Proposal for a Detection Method Using Data Size of SSH Password Cracking Attacks and its Operational Evaluations

Kouji Shimizu[1,a)] Tomoya Kotone[1,†1] Minoru Ikebe[2] Kazuyuki Yoshida[3]

Received: June 27, 2016, Accepted: December 1, 2016

Abstract: There are many malicious attacks in the Internet. In particular, there are many illegal accesses into SSH servers. So, we have been developing a SSH Password Cracking Attack Detection system (called SCRAD) in order to detection for SSH password cracking attacks. Our system detects attacker's connection using the number of packets per connection between a SSH client and server. We analyzed the operational results of SCRAD system. We found some false positives. The cause of false positives was a small amount of data communication using scp or rsync commands by normal users. Therefore, we investigated the connection of scp's communication. As a result, the number of packets is similar to the scp and attacker's connection. However, data size in the scp connection is larger than data size in the attacker's connection. In this paper, we propose a new detection method using data size to avoid the false positives. We compared the packet counts based method and the data size based method. In the experimental results, the attacker detection rate was the same of the two methods. However, the normal user discrimination rate of the data size based method was 22.1% higher than the packet counts based method. We confirmed that the proposed method is effective for the SSH password cracking attacks. In addition, we are operating the SCRAD system with data size based detection method. We report the operational results of SCRAD from February 2015 to May 2016. The SCRAD was detected 99.7% of attacker's connections. On the other hand, normal user connection discrimination rate was 72.2%. The operational results show that SCRAD reduced the risk of penetrated SSH servers, and SCRAD occured some false positives.

Keywords: SSH, password cracking attack, TCP, anomaly based IDS

1. はじめに

インターネットの普及にともない，我々はネットワークを通して様々な情報のやりとりをしている．そのため現在では，インターネットは社会的基盤の 1 つとして生活に不可欠な存在である．しかし，インターネットを利用した不正通信も数多く存在する．特に，不正アクセス行為の発生件数は増加傾向にあり，不正アクセス行為後の被害としては，インターネットバンキングの不正送金や，オンラインショッピングのアカウント乗っ取りなど生活に影響を与えるものが数多くある．不正アクセスは，ソフトウェアの脆弱性や脆弱なパスワードを利用し，サーバへ侵入する．サーバへ侵入するために，攻撃者は SSH サーバに対するパスワードクラッキング攻撃を仕掛ける．JPCERT/CC の報告では SSH に対する scan 攻撃が観測されている [1]．JPCERT/CC の最新の報告 [2] では 2016 年 1〜3 月における SSH に対する攻撃は，宛先ポート番号の上位 5 件からはなくなったものの，JPCERT/CC が運用しているインターネット定点観測システムの TSUBAME の SSH ポートに対する観測結果は以前と変わらないアクセス数で推移している [3]．また，大分大学宛のポート別 SYN パケット数の調査結果を**表 1** に示す．SSH に対する SYN パケット数の順位は以前と比べ下がっているが，SYN パケット数は約 2 年前は 1 日あたり 1,831,496 パケットであり，大きな変化はないため，継続的に SSH パスワードクラッキング攻撃が行われている．

我々は SSH へのパスワードクラッキング攻撃を検知することを目的とした「SSH パスワードクラッキング攻撃検知システム（SCRAD）」[4] を開発・運用してきた．本システムは送信元とサーバの間で送受信する 1 コネクションあたりのパケット数を計数し，パスワードクラッキング攻撃を検知する．しかし，これまでの SCRAD の検証結果から，パケット数を用いる検知手法では，scp や rsync で少量のデータを送受信する際に正規ユーザを攻撃者として誤検知することが判明している [5]．以降では，正規ユーザを攻撃者として誤検知することを正規ユーザの誤検知と表現する．そこで本論文では，1 コネクションあたりのパケット数を用いた検知手法と同等の検知率を実現し，正規ユーザの誤検知を改善することを目的として，データサイズを用いる検知手法を提案する．提案手法における検知率と，従来のパケット数による検知手法における検知率を比較して，提案手法の有用性について評価した．

本論文の構成を以下に示す．2 章では，SSH パスワードクラッキング攻撃検知に関する研究について述べる．3 章では，SSH パスワードクラッキング攻撃検知システム（SCRAD）について述べる．4 章では，1 コネクションあたりのデータサイズに着目した検知手法を提案し，評価する．5 章では，2016 年 5 月 31 日までの運用結果から SCRAD の検知精度について述べる．6 章では，まとめと今後の課題について述べる．

表 1 大分大学宛のポート別 SYN パケット数（収集期間：2016 年 4 月 6 日 18:50〜2016 年 4 月 7 日 19:40）
Table 1 Number of SYN packets to TCP ports for Oita University's network (duration: April 6, 2016, 18:50 JST ∼ April 7, 2016, 19:40 JST).

ポート番号	SYN パケット数	用途
50382	18,461,323	用途不明
50390	18,447,508	用途不明
23	16,047,094	Telnet
80	8,126,476	HTTP
445	5,244,033	Microsoft Directory Service
443	4,171,501	HTTPS
22	1,863,151	SSH

2. SSH パスワードクラッキング攻撃検知に関する研究

2.1 SSH パスワードクラッキング攻撃検知システムにおける要件

我々は，SSH パスワードクラッキング攻撃検知システムを開発する際に，パスワードクラッキングによる SSH サーバに対する侵入の脅威や，学内端末のボット感染などを考慮し，以下の 3 つの用件を定義した．

(1) リアルタイム検知可能であること
(2) 攻撃者 IP アドレスの遮断が可能であること
(3) 学外からの攻撃だけでなく，学内のボット感染ホストの検知が可能であること

2.2 関連研究

本節では，SSH パスワードクラッキング攻撃の検知に関する関連研究を紹介する．SSH パスワードクラッキング攻撃を検知する手法として，ネットワークトラフィックベースの手法とホストの認証ログベースの手法がある．

トラフィックの解析から SSH サーバへのパスワードクラッキング攻撃を検知する手法として Vykopal らの手法 [6] があげられる．Vykopal らは，SSH へのパスワードクラッキング攻撃を分析し，辞書攻撃によるトラフィックのパ

[1] 大分大学大学院工学研究科知能情報システム工学専攻
Course of Computer Science and Intelligent Systems, Graduate School of Engineering, Oita University, Oita 870–1192, Japan
[2] 大分大学工学部知能情報システム工学科
Department of Computer Science and Intelligent Systems, Faculty of Engineering, Oita University, Oita 870–1192, Japan
[3] 大分大学学術情報拠点情報基盤センター
Center for Academic Information and Library Services, Oita University, Oita 870–1192, Japan
[†1] 現在，三菱電機株式会社
Presently with Mitsubishi Electric corporation
[a)] v15e3014@oita-u.ac.jp

ターンを一般的な SSH トラフィックと比較することでリアルタイムに SSH に対するパスワードクラッキング攻撃を検知するとともに攻撃成功の判別を可能にした．SSH パスワードクラッキング攻撃の検知および SSH パスワードクラッキング攻撃の成功の判別には，決定木手法を用いている．この決定木にはトラフィックにおけるパケットの送信間隔や送信回数，データサイズの情報が含まれている．この手法の問題点は，SSH サーバ監視のために Nagios [7] によって生成された SSH のトラフィックパターンが攻撃パターンと類似するため誤検知するという点がある．

ホストの認証ログベースの手法として大隅らの手法 [8] があげられる．大隅らは，SSH サーバのアクセスログを監視することにより，パスワード認証の SSH サーバに対する総当たり攻撃や辞書攻撃を検知し，不正なアクセスを動的に拒否する方式を提案した．彼らの手法は，組織内の各 SSH サーバのアクセスログを syslog サーバへ集約し，パスワード認証に失敗したログを抽出する．syslog サーバでは，単位時間あたりのパスワード認証エラー回数を集計し，しきい値が超過した送信元を攻撃者として検知する．攻撃者を検知後，syslog サーバは攻撃者の IP アドレスを組織内の各 SSH サーバに通知し，各 SSH サーバで攻撃者との接続を拒否する．この手法は，組織内の全 SSH サーバのアクセスログを監視し，攻撃を検知する．よって，インターネットから組織内への全 SSH サーバへの攻撃を防ぐことができる．しかし，組織内のボットに感染したホストが組織内からインターネットへパスワードクラッキング攻撃を仕掛けている場合，そのホストを発見することが困難であるという問題点がある．

3. SSH パスワードクラッキング攻撃検知（SCRAD）システム

本章では 2.1 節で述べた要件を満たす SSH パスワードクラッキング攻撃検知（SCRAD: SSH password CRacking Attack Detection）システムのシステム構成や従来手法である 1 コネクションに送受信するパケット数を用いる検知手法における問題点について述べる．

3.1 システム概要

SCRAD は，インターネットと学内ネットワークの境界を通過するパケットを LAN スイッチのポートミラー機能で複製し，TCP/22 番ポートに関するパケットを tcpdump のフィルタ機能を用いて抽出する（**図 1**）．送信元 IP アドレスをキーとして，各コネクションの確立から終了までのパケット送受信回数（PN_{all}）を監視する．さらに送信元 IP アドレスごとの PN_{all} がしきい値以下のコネクション数を計数することで，リアルタイムにパスワードクラッキング攻撃を検知する．検知した送信元 IP アドレスは送信元 IP アドレスのツリーから攻撃者 IP アドレスのツリーへ移す．攻撃者 IP アドレスツリーに挿入した IP アドレスは経路制御で SSH サーバからの戻りのパケットを破棄している [9]．SCRAD システムは，インターネットと学内ネットワークの間を流れる双方向のパケットを入力とするため，送信元が学内外のどちらの場合でもパスワードクラッキング攻撃を検知できる．SCRAD システムには，tcpdump を用いてリアルタイムにパケットをキャプチャして，攻撃検知するオンラインモードと，あらかじめ収集した pcap ファイルを用いて，事後において攻撃を判定するオフラインモードが存在する．

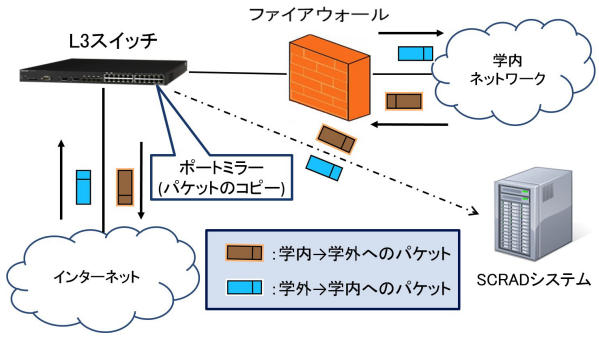


図 1 SCRAD システムの構成図
Fig. 1 Overview of SCRAD system.

3.2 パケット数を用いる検知手法（SCRAD-CP）

我々の先行研究 [5] において，SSH クライアントとサーバ間の 1 コネクションあたりのパケット数を調査した．調査結果から，SSH パスワードクラッキング攻撃を検知するためのしきい値は 1 コネクションあたりのパケット送受信回数 PN_{all} において 50 パケット未満と定義した．SCRAD システムにおいて，PN_{all} とは，送信元の SYN パケットを観測後，送信元と SSH サーバ間で最初の FIN パケットまたは RST パケットを観測するまでのパケット数である．また先行研究 [4], [5] により，パケット数の計数からデータサイズ 0 の TCP パケット（ACK パケット）と再送パケットを除外している．これは，データサイズ 0 の TCP パケットや再送パケットによって PN_{all} がしきい値を超えることがあるためである．PN_{all} から，データサイズ 0 の TCP パケット，再送パケットを除いた本手法で用いる検知のためのパケット送受信回数を $PN_{detection}$ とする．また，$PN_{detection}$ がしきい値（50 パケット）未満のコネクションを FAIL コネクション，しきい値以上のコネクションを SUCCESS コネクションと定義する．また，正規ユーザの誤検知を防ぐため，10 回連続で FAIL コネクションを観測した送信元を攻撃者として検知する．本論文では，本節で述べた 1 コネクションに送受信するパケット数を計数する検知手法を SCRAD-CP（Count Packets）と呼ぶ．SCRAD-CP の攻撃者検知基準を以下に示す．

- $PN_{detection} = PN_{all}$ − データサイズが 0 の TCP パ

```
コマンド：scp ./file server1:
31 133.37.X -> 133.37.Y 2014 12/08 19:27:17
コマンド：scp server1:file1 .
26 133.37.X -> 133.37.Y 2014 12/08 19:31:00
```

図 2 scp によるコネクションの調査結果
Fig. 2 Results of SCP command's connections.

ケット数 − 再送パケット数

- $PN_{detection}$ が 50 パケット未満ならば FAIL コネクションと判定
- FAIL コネクションを 10 回連続で観測した送信元を攻撃者として検知

3.3 SCRAD-CP における問題点

SCRAD-CP では，まれに正規ユーザの誤検知が発生していた．誤検知した IP アドレスのユーザにヒアリングし原因を調査したところ，誤検知の原因と考えられる 2 つの要素があった．(1) ssh-agent に，ssh-add コマンドを用いて秘密鍵とパスフレーズの組合せを記憶させた公開鍵認証による自動ログイン機能を用いており，(2) ファイルを送受信する際に，scp コマンドや rsync コマンドにより少量のデータを送受信する SSH コネクションを誤検知していた．そこで，実際に scp コマンドによりファイルを送受信するコネクションで流れるパケットを tcpdump で収集し，収集した pcap ファイルを SCRAD-CP に入力することにより scp によるコネクションを調査した．調査したコネクションのログを**図 2** に示す．図 2 のログの一番左のフィールドがパケット数を示している．調査結果より，scp によるコネクションにおけるパケット数が 50 パケット未満となることが判明した．SSH パスワードクラッキング攻撃では，パスワード認証試行後，一定回数認証に失敗するとセッションが終了する．しかし，scp によるコネクションにおいては，ユーザ認証後にデータの送受信が生じる．そこで，データサイズに着目すると，ユーザ認証後にデータを送受信する scp によるコネクションとデータの送受信がないまま終了した FAIL コネクションを判別できると予測した．

4. データサイズによる検知手法（SCRAD-SS）の提案と評価

我々は，正規ユーザと攻撃者における SSH コネクションで生じるデータサイズの差異に着目した新たな検知手法を提案する．提案手法では，SCRAD-CP と同等の検知精度を維持しつつ，SCRAD-CP で生じた正規ユーザの誤検知を改善することを目的としている．予備調査として，正規ユーザと攻撃者の通信を判別するため，正規ユーザのコネクションと攻撃者のコネクションにおいて送受信されるパケットのデータサイズを調査した．なお，データサイズの算出には，クライアントとサーバの双方向におけるコネクションの SYN パケットと FIN または RST パケットのシーケンス番号の差を用いている．また，調査には，クライアントとサーバの双方向におけるデータサイズを合計したデータサイズを用いる．データサイズの算出にシーケンス番号の差を用いるため，提案手法を SCRAD-SS（SCRAD-Subtract Sequence number）と呼ぶ．調査結果をもとに攻撃者の検知のためのしきい値を設定する．

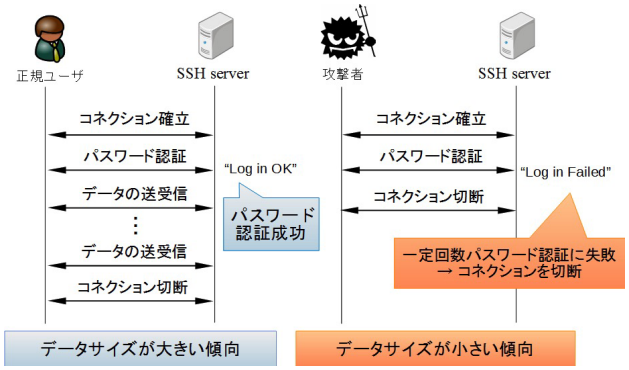


図 3 1 コネクションにおけるデータサイズの違い
Fig. 3 Differences of data size between normal user and attacker connections.

4.1 SSH コネクションにおける送受信するデータサイズの違い

正規ユーザと攻撃者の SSH コネクションのデータサイズの違いを**図 3** に示す．正規ユーザと SSH サーバの通信は，まずユーザ認証プロセスによりユーザを認証する．その後にデータを送受信する．よって，正規ユーザによる 1 コネクションあたりに送受信するデータサイズは大きくなる傾向にある．一方，攻撃者と SSH サーバの通信は，ブルートフォース攻撃や辞書攻撃によって何度もユーザ認証を繰り返す．一定回数（通常は 3 回）パスワード認証に失敗した場合，TCP コネクションは切断される．そのため，正規ユーザの通信にあるデータの送受信は生じない．よって，攻撃者による 1 コネクションあたりのデータサイズは小さくなる傾向にある．

4.2 SSH コネクションにおけるデータサイズの予備調査

パスワードクラッキング攻撃を仕掛けた通信をデータサイズによって検知するためのしきい値を設定するため，攻撃者のコネクションで送受信するデータサイズと正規ユーザのコネクションで送受信するデータサイズを調査した．調査内容は以下の 2 点である．

- 正規ユーザが送受信するデータサイズの最小値
- 攻撃者が送受信するデータサイズの最大値

調査方法については，それぞれの節で詳しく述べる．

表 2 データサイズの調査結果（各 3 回の平均値）
Table 2 Results of data size in each connection (Each value is an average bytes of 3 times).

調査コネクション	clt[*1]→ srv	srv[*2]→ clt	合計
何もせず切断	3,835	2,829	6,664
scp(送信)	3,835	2,877	6,712
scp(受信)	4,011	2,941	6,952
rsync	4,187	3,053	7,240

[*1] clt：SSH クライアント
[*2] srv：SSH サーバ

4.2.1 正規ユーザのコネクションにおけるデータサイズの調査

まずは，正規ユーザのコネクションで送受信するデータサイズの最小値を求めるため，以下の 3 つの SSH コネクションのパケットを tcpdump で収集し調査した．本論文では，公開鍵認証方式を用いた場合のコネクションを調査した．また，調査の試行回数は 3 回であり，数値は 3 回の平均値である．データサイズの調査結果を**表 2** に示す．

(1) SSH サーバにログインして何もせずにコネクションを切断
実際に SSH サーバへログインして何もせずにコネクションを切断した通信（”ssh srv :” を実行）において，送受信したパケット数は 24 パケットであり，データサイズの合計は 6,664 バイトであった．

(2) scp コマンドで 0 バイトのファイルを送受信
SSH 経由でファイルをコピーする scp コマンドで，クライアントからサーバへファイルを送信，クライアントでサーバからファイルを受信の 2 つのパターンで実験した．学内のある SSH サーバから 0 バイトのファイルを scp コマンドを用いて受信したコネクションにおいて送受信されたパケット数の合計は 31 であり，データサイズの合計は 6,952 バイトであった．同じ SSH サーバへ 0 バイトのファイルを送信したところ，送受信したパケット数は 32 パケットであり，データサイズの合計は 6,712 バイトであった．

(3) rsync コマンドで互いに同じファイルを同期
ファイルを同期する rsync コマンドを，SSH 経由で利用し，差分のないファイルを同期したときに送受信されたパケット数は 34 パケットであり，データサイズの合計は 7,240 バイトであった．

よって，今回計測した 3 つのコネクションにおける最小値は (1) 何もせずに切断した場合の 6,664 バイトであった（表 2）．

4.2.2 攻撃者のコネクションにおけるデータサイズの調査

次に，攻撃者のコネクションで送受信するデータサイズの最大値を調査した．2014 年 5 月 1 日から 2014 年 5 月 31 日の 1 カ月間に図 1 の L3 スイッチから収集したパケットデータを用いて，1 コネクションあたりの送信元と SSH サーバの間で送受信されるデータサイズを調査した．SCRAD-CP と SCRAD-SS では，正規ユーザと攻撃者の 2 つの分類であるが，正規ユーザの誤検知や，攻撃者の検知漏れを含む可能性がある．そこで，より正確な攻撃者のコネクションにおけるデータサイズの最大値を求めるために，パケットデータに含まれる SSH コネクションを接続した送信元をすべて調査し，正規ユーザ・攻撃者・不明なユーザの 3 つに分類した．分類した送信元をそれぞれの送信元における真の値と定義する．送信元の分類方法を以下に示す．

- 正規ユーザ
  $PN_{detection}$ が 100 パケット以上のコネクションを 1 回以上検知した送信元
- 不明なユーザ
  正規ユーザ以外の送信元であり，かつ検知したコネクション数が 3 回以下
- 攻撃者
  正規ユーザ，不明なユーザ以外の送信元

正規ユーザは，1 コネクションあたりのパケット数が 100 パケット以上のコネクションを 1 回以上検知した送信元である．これまでの SCRAD-CP の運用結果により，100 パケット以上のコネクションの場合は，パスワードクラッキング攻撃の通信ではないことが判明している．よって上記の挙動を示す送信元は正規ユーザと定義した．不明なユーザとは，100 パケット未満のコネクションを 3 回以下しか検知しなかった送信元である．このような送信元は，判定に用いるコネクション数が少ないため，正規ユーザや攻撃者とも判断できない．よって上記の送信元を不明なユーザと定義した．SCRAD においては，不明なユーザは SUCCESS コネクションをつながず，かつ攻撃者として検知できる回数（10 回）まで FAIL コネクションをつないでいないユーザであることから，攻撃者として検知できないユーザである．攻撃者は，正規ユーザにも不明なユーザにも属さない送信元である．真の値を調査した分類結果を**表 3** に示す．表 3 の不明なユーザについては，確実に攻撃者，もしくは正規ユーザと判断できる送信元ではないため，今回の調査からは除外した．不明なユーザに関しては，我々の調査 [10] において詳細に分析している．表 3 の分類に基づいて，真の正規ユーザと真の攻撃者によるコネクションのデータサイズを**図 4** に示す．図 4 から，真の正規ユーザにおけるコネクションのデータサイズはばらつきが大きいことが分かる．一方，真の攻撃者によるコネクションのデータサイズは，10,000 バイト以下に集まっており，ばらつきが小さい．そこで，データサイズ 0〜10,000 バイトまでの範囲を抽出したグラフを**図 5** に示す．10,000 バイト以下の範囲を調査したところ攻撃者のコネクションで送受信するデータサイズの最大値は 4,539 バイトであった．図 5 より，正規ユーザの送受信するデータサイズが攻

表 3 2014 年 5 月における真の値
Table 3 True values in May 2014.

(a) SSH クライアント

クライアントの種類	真の値
正規ユーザ	59
攻撃者	202
不明なユーザ	47
合計	308

(b) SSH コネクション

通信の種類	真の値
正規ユーザによるコネクション	1,206
攻撃者によるコネクション	7,258
不明なユーザによるコネクション	61
合計	8,525

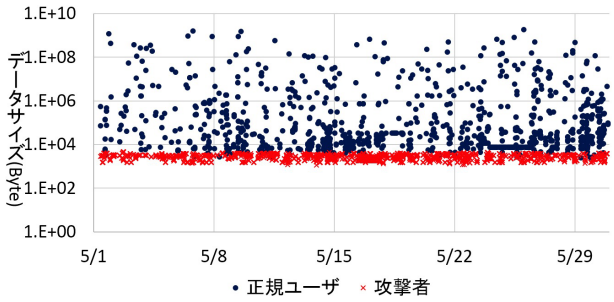


図 4 真の値におけるデータサイズ
Fig. 4 Data size of true normal user and attacker.

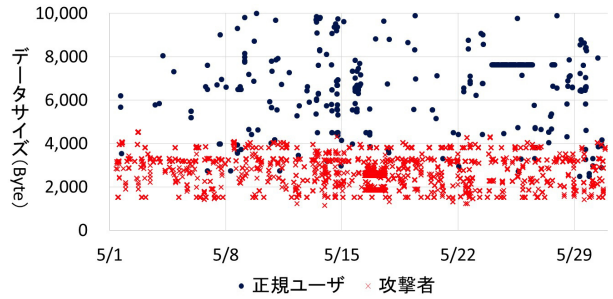


図 5 真の値におけるデータサイズ（0〜10,000 バイト）
Fig. 5 Data size of true normal user and attacker (0 ∼ 10,000 bytes).

撃者の送受信するデータサイズと類似するコネクションを観測した．このようなコネクションは，正規ユーザがパスワード入力を失敗したことで，図 3 にある攻撃者に類似したコネクションになったためと考えられる．

4.3 データサイズによる検知手法（SCRAD-SS）における検知基準

4.2 節の調査結果から，SCRAD-SS における検知基準を以下のように定めた．

- 1 コネクションで送受信するデータサイズが 5,000 バイト未満のコネクションを FAIL コネクションとする．
- FAIL コネクションを 10 回連続で検知した送信元を攻撃者として検知する．

表 4 SCRAD-CP と SCRAD-SS の比較
Table 4 Comparison between SCRAD-CP and SCRAD-SS.

(a) SSH クライアント検知数

クライアントの種類	SCRAD-CP	SCRAD-SS
正規ユーザ	66	70
攻撃者	156	155
重複検知	(4)	(4)
未分類*	167	164
合計**	385	385

* 未分類は，正規ユーザにも攻撃者にも分類されなかった SSH クライアント
** 合計 = 攻撃者数 + 正規ユーザ数 − 重複検知数 + 未分類数

(b) SSH コネクション判定数

通信の種類	SCRAD-CP	SCRAD-SS
SUCCESS コネクション	551	741
FAIL コネクション	5,253	5,063
合計	5,804	5,804

4.4 SCRAD-CP と SCRAD-SS における検知精度の比較

4.4.1 SCRAD-CP と SCRAD-SS の実験結果

従来までのパケット数を用いる検知手法と，提案手法であるデータサイズを用いる検知手法を比較した．比較に用いたデータセットは，2014 年 7 月 1 日から 2014 年 7 月 31 日の期間に収集したパケットデータである．それぞれの検知手法による検知結果のうち，**表 4** (a) に SSH クライアント検知数を，表 4 (b) にコネクション判定数を示す．表 4 (b) は，4.4.5 項の検知精度の比較で用いる．

4.4.2 SSH クライアントの検知数の比較

表 4 (a) における SSH クライアント検知数を比較する．表 4 (a) において重複検知した SSH クライアントは，攻撃者，正規ユーザの両方の判定をされた SSH クライアントであり，SSH サーバへの不正侵入に成功した攻撃者か，誤検知した正規ユーザの可能性がある．表 4 (a) において，SCRAD-CP と SCRAD-SS で検知クライアント数に差が生じていた．検知クライアント数の差は SCRAD-CP においてパケット送受信回数が 50 パケット未満の FAIL コネクションを，SCRAD-SS では同一の SSH クライアントのコネクションをデータサイズ 1,000 バイト以上の SUCCESS コネクションとして判定し，正規ユーザに分類したために生じた．SCRAD-CP，SCRAD-SS とも，FAIL コネクションを 10 回未満観測した SSH クライアントを，正規ユーザとも攻撃者としても判断できない未分類の SSH クライアントとして，表 4 (a) に記載している．未分類となった SSH

```
4349 16 （ア） -> （オ） F 2014 07/06 07:20:51
4547 16 （ア） -> （カ） F 2014 07/06 07:21:17
4349 16 （ア） -> （キ） F 2014 07/06 07:21:21
6195 24 （イ） -> （ク） F 2014 07/05 01:05:36
5441 21 （ア） -> （ケ） S 2014 07/06 07:20:54
9098 18 （ウ） -> （コ） S 2014 07/15 09:02:39
2850499900 9 （エ） -> （サ） S 2014 07/29 00:32:56
```

図 6 SCRAD-SS で正規ユーザに分類されたユーザによるコネクション

Fig. 6 Connections by users classified as normal user in SCRAD-SS.

クライアントの詳細については別稿 [10] において報告している．

SCRAD-CP で検知した正規ユーザ 66 件は SCRAD-SS で検知した正規ユーザ 70 件にすべて含まれていた．これらの分類の妥当性について検証するために，SCRAD-SS でのみ正規ユーザに分類した 4 件の送信元 IP アドレスによるコネクションについて調査した．SCRAD-SS においてのみ正規ユーザと分類した 4 件の送信元 IP アドレスによるコネクションを**図 6** に示す．S は SUCCESS コネクション，F は FAIL コネクションを示す．図 6 における送信元のうち，学内の IP アドレスを持つ送信元（ア）は，連続して FAIL コネクションを接続する挙動があり，SUCCESS コネクションは 5,000 バイトをわずかに超過した 1 件のみ観測している．そのため，送信元（ア）は，学内のボット感染ホストの疑いがあり，SCRAD-SS が攻撃者を検知漏れした可能性がある．また，送信元（イ）と送信元（ウ）に関しては，外部の IP アドレスであり，図 6 に示す時間以外にはコネクションを観測していない．この 2 件の送信元 IP アドレスによるコネクションはいずれもパケット数が 50 パケット以下であり，かつデータサイズが 5,000 バイトを超えているため，ファイルの送受信などのコマンドによりデータのやりとりをした正規ユーザと考えられる．外部の送信元 IP アドレス送信元（エ）に関しては，9 パケットで 2,850,499,900 バイトのデータを送受信することは，MSS（Maximum Segment Size）の観点から考えられないため，SCRAD-SS のデータサイズ計算方法に問題があると考えられる．比較に用いたデータセットより，送信元（エ）のコネクションに該当するパケットを抽出したところ，同一の送信元ポート番号を用いて複数のコネクションを接続していた．さらに，1 つのコネクションの終了時に FYN パケットを送信せずに次のコネクションの接続を要求する SYN パケットを送信していた．実際に 1 コネクションで送受信したデータサイズは，パケットのシーケンス番号より，22 バイトであった．さらに，TCP コネクション確立後にまったくパケットを送信しないという挙動から，送信元（エ）は scan 攻撃を仕掛けていたと考えられる．SCRAD-CP と SCRAD-SS では，scan 攻撃の場合に，$PN_{detection}$ が 5 パケット以下をしきい値としてコネクション管理しないように設計しているが，同一の送信元ポート番号を使われたことにより，複数のコネクションを同一のコネクションと判断したため，1 コネクションにおけるパケット送受信回数 $PN_{detection}$ が scan 攻撃と判断するしきい値を超えた．そのため SCRAD-CP と SCRAD-SS が scan 攻撃としてコネクション情報を削除できなかったと考えられる．SCRAD-CP と SCRAD-SS では，送信元 IP アドレスと送信元ポート番号，宛先 IP アドレス，宛先ポート番号によりコネクションを管理しているため，同一の送信元ポート番号を用いてコネクションを接続し，FYN パケットを送信せずに SYN パケットを受信した場合，SCRAD-CP と SCRAD-SS のコネクション管理における状態遷移が FYN パケット待ちの状態のままになる．次のコネクションにおける FYN パケットのシーケンス番号を用いてデータサイズを計算することになるため，データサイズの値が実際とは違う値になったと考えられる．

4.4.3 誤検知の検証

4.2 節の分類方法に基づいて，2014 年 7 月 1 日から 7 月 31 日における真の値を調査した．

表 4 (a) に示した攻撃者の検知結果より，SCRAD-CP では 2 件，SCRAD-SS では 1 件の誤検知があった．SCRAD-CP，SCRAD-SS では同一の SSH クライアントを誤検知していた．この SSH クライアントについては 4.4.6 項の重複検知した SSH クライアントの分析において述べる．SCRAD-CP における誤検知 2 件のうち，1 件は SCRAD-SS と共通な SSH クライアントであった．もう 1 件の誤検知した SSH クライアントはパケット送受信回数（$PN_{detection}$）38，データサイズ 8,473 バイトのコネクションを 10 回連続で観測していた．SCRAD-CP により誤検知していた SSH クライアントは SCRAD-SS ではデータサイズが 5,000 バイトを超えているため，正規ユーザと正しく分類できていた．

4.4.4 検知漏れの検証

次に，検知漏れを検証した．調査結果を**表 5** に示す．表 5 (a) における真の攻撃者 172 件のうち，SCRAD-CP，SCRAD-SS ともに 154 件の真の攻撃者を攻撃者として検知した．いずれの手法においても 18 件は攻撃者と分類できていなかった．どちらの検知手法においても，真の攻撃者として分類された SSH クライアント 172 件のうち 18 件の SSH クライアントを攻撃者として検知できなかった．これらの SSH クライアントは FAIL コネクションを 4 から 9 回連続で観測しており，攻撃者の可能性がある．2 つの手法における攻撃者クライアントの検知漏れ率を以下の式より求めた．

表 5 2014 年 7 月における真の値
Table 5 True values in July 2014.

(a) SSH クライアント

クライアントの種類	真の値
真の正規ユーザ	63
真の攻撃者	172
不明なユーザ	150
合計	385

(b) SSH コネクション

通信の種類	真の値
真の正規ユーザによるコネクション	803
真の攻撃者によるコネクション	4,821
不明なユーザによるコネクション	180
合計	5,804

検知漏れ率

$$= \frac{\text{検知システムが検知できなかった真の攻撃者数}}{\text{真の攻撃者数}}$$

$$\text{検知漏れ率 (SCRAD-CP)} = \frac{18}{172} = 0.104$$

$$\text{検知漏れ率 (SCRAD-SS)} = \frac{18}{172} = 0.104$$

上記の結果より，2 つの手法において，検知漏れ率に差異はなかった．

誤検知と検知漏れはトレードオフの関係にある．我々は，正規ユーザの誤検知を防止することを重要視している．SCRAD-CP，SCRAD-SS において FAIL コネクションを 10 回連続で観測した送信元 IP アドレスを攻撃者として検知することで，誤検知を防止している．そのため，本項で示した検知漏れしていた攻撃者は FAIL コネクションを 4 から 9 回連続観測しており，我々の設定したしきい値以下であったため，攻撃者を検知漏れした．

4.4.5 検知精度の比較

真の値による分類と SCRAD-CP，SCRAD-SS におけるコネクションの判定結果の関係を**表 6** に示す．たとえば，表 6 における A とは，真の値における分類では真の正規ユーザに分類された送信元によるコネクションであり，かつ SCRAD-CP と SCRAD-SS で SUCCESS コネクションに分類されたコネクションを示している．**表 7** は，SCRAD-CP と SCRAD-SS それぞれにおいて判定した SUCCESS コネクションと FAIL コネクションの分類が真の値における正規コネクション，攻撃コネクションの分類と一致した値を示す．

表 7 中の値には，表 4 (b) のコネクション判定結果と値が一致しないものがある．たとえば，表 7 (a) における SCRAD-CP にて FAIL コネクションに分類したコネクションと，真の値により正規ユーザに分類した SSH クライアントによるコネクションと一致した 262 件に，SCRAD-CP にて FAIL コネクションに分類したコネクションと，真の値により攻撃者に分類した SSH クライアントによるコネクションと一致した 4,811 件を加算した 5,073 件は，表 4 (b) 中の SCRAD-CP が判定した FAIL コネクションの総数 5,252 件と一致しない．これは，真の値により不明なユーザに分類した SSH クライアントによる FAIL コネクション数が，表 4 (b) には含まれているが，表 7 (a) には真の値により攻撃者，もしくは正規ユーザに分類した SSH クライアントによるコネクションしか含まれていないためである．また，表 7 (b) における SUCCESS コネクションの判定となった計 735 件は，表 4 (b) における SCRAD-SS の SUCCESS コネクション判定数 741 件と異なる．これは真の値により不明なユーザに分類した SSH クライアントによる $PN_{detection}$ が 50 パケット未満，かつ 1 コネクションに送受信するデータサイズが 5,000 バイト以上のコネクションの接続が存在したため，SCRAD-SS が SUCCESS コネクションと判定したコネクションの中に，真の値により攻撃者とも正規ユーザとも分類されなかった SSH クライアントによるコネクションが存在しているためである．

真の値に基づいて，SCRAD-CP と SCRAD-SS におけるコネクションを真の攻撃者のコネクションと真の正規ユーザのコネクションに分け，それぞれの検知手法による検知精度を評価した．検知精度の評価には，表 6 の分類を用いて正規ユーザのコネクションと攻撃者のコネクションそれぞれの判定率を求めた．ここで用いる判定率の計算式は以下のとおりである．

真の正規ユーザコネクション判定率

表 6 SCRAD システムにおけるコネクションの判定結果と真の値による分類における関係
Table 6 Relationship between results of SCRAD and classification of true values.

真の値 ＼ 判定結果	SUCCESS コネクション	FAIL コネクション
正規コネクション	A	B
攻撃コネクション	C	D

表 7 真の値による分類とそれぞれの手法による検知結果の関係
Table 7 Relationship between classification by true values and detection results of two method.

(a) SCRAD-CP の判定結果と真の値の関係

真の値 ＼ 判定結果	SUCCESS コネクション	FAIL コネクション
正規コネクション	541	262
攻撃コネクション	10	4,811

(b) SCRAD-SS の判定結果と真の値の関係

真の値 ＼ 判定結果	SUCCESS コネクション	FAIL コネクション
正規コネクション	724	79
攻撃コネクション	11	4,810

表 8 コネクション別の判定精度の比較
Table 8 Comparison of detection accuracy by connections.

コネクション	SCRAD-CP(%)	SCRAD-SS(%)
攻撃者	99.7	99.7
正規ユーザ	67.3	90.1

$$\frac{A}{A+B}$$

真の攻撃者コネクション判定率

$$\frac{D}{C+D}$$

上記の判定率の計算式と表 7 に示した SCRAD-CP, SCRAD-SS の求めた検知精度を**表 8** に示す．SCRAD-CP では，正しく判別できた正規ユーザのコネクションは 67.3%であったが，SCRAD-SS では 90.1%の正規ユーザのコネクションを SUCCESS コネクションとして検知できた．よって，SCRAD-CP と比べ，SCRAD-SS の方がより正確に正規ユーザのコネクションを判定できた．

4.4.6 重複検知

重複検知とは，SCRAD-CP および SCRAD-SS において，攻撃者の検知条件を満たし，かつ SUCCESS コネクションを観測することである．どちらの検知手法においても，4 つの SSH クライアントを重複検知した．この 4 つの送信元を（ア），（ス），（セ），（ソ）とする．この 4 つの送信元の通信ログを**図 7** に示す．（ア）の通信ログを図 7 (a) に示す．図 7 (a) における学内の IP アドレスを持つ送信元（ア）は，4.4.2 項で述べた検知漏れの可能性がある SSH クライアントである．外部の特定の宛先 IP アドレスへ不定期に通信をしていた（図 7 (a)）が，一方で 4.4.2 項で述べた挙動を考慮すると，学内のボット感染ホストである可能性がある．

送信元（ス）による通信ログを図 7 (b) に示す．学内の IP アドレスを持つ送信元（ス）は他大学のコンピュータシステムのサーバとのコネクションである．よって，この送信元（ス）の送信元に関しては，攻撃とは考えにくいため，誤検知したと考えられる．なお，送信元（ス）は，4.4.3 項で述べたように，SCRAD-CP と SCRAD-SS 両方の手法で誤検知した SSH クライアントである．

（セ），（ソ）の通信ログを図 7 (c) に示す．学外の送信元（セ），（ソ）の通信は，多くの宛先とコネクションをつないでいた．（図 7 (c)）．また，短期間に大量のコネクションをつなぐ挙動を示している．よって，送信元（セ），（ソ）は攻撃者と考えられ，SUCCESS コネクションをつながれた宛先 SSH サーバに侵入された可能性がある．

4.4.7 検知精度の比較結果のまとめ

本節で示した SCRAD-SS と SCRAD-CP における検知精度の比較の結果，表 8 より SCRAD-SS では SCRAD-CP と同等の攻撃者検知率であり，より正確に正規ユーザのコネクションを分類できていた．そのため，我々は，攻撃者検知において，SCRAD-CP よりも SCRAD-SS が有用であると判断した．

```
28933 65 （ア） -> （タ） S 2014 07/03 16:05:08
18580 191 （ア） -> （チ） S 2014 07/03 16:09:13
10736445 7865 （ア） -> （タ） S 2014 07/03 17:18:15
224745 358 （ア） -> （チ） S 2014 07/03 18:07:42
1153129 1127 （ア） -> （タ） S 2014 07/04 18:05:47
2627 14 （ア） -> （チ） F 2014 07/24 13:39:16
2627 14 （ア） -> （チ） F 2014 07/24 13:39:17
2627 14 （ア） -> （チ） F 2014 07/24 13:39:18
2627 14 （ア） -> （タ） F 2014 07/30 14:04:56
```

(a) 学内のボット感染ホストと考えられる SSH クライアントの通信ログ

```
2627 14 （ス） -> （ツ） 2014 07/24 13:39:09
2627 14 （ス） -> （ツ） 2014 07/24 13:39:14
2627 14 （ス） -> （ツ） 2014 07/24 13:39:15
2627 14 （ス） -> （ツ） 2014 07/24 13:39:15
2627 14 （ス） -> （ツ） 2014 07/24 13:39:16
2627 14 （ス） -> （ツ） 2014 07/24 13:39:17
13799961 10062 （ス） -> （ツ） 2014 07/22 11:23:55
4008633 4526 （ス） -> （テ） 2014 07/22 14:03:07
35807 218 （ス） -> （テ） 2014 07/24 17:41:45
253779 829 （ス） -> （ツ） 2014 07/30 14:04:00
```

(b) 誤検知したと考えられる SSH クライアントの通信ログ

```
4746 23 （セ） -> （ト） F 2014 07/01 09:27:15
2542 12 （セ） -> （ナ） F 2014 07/15 01:23:17
1895 15 （セ） -> （ニ） F 2014 07/15 01:26:39
2925 13 （セ） -> （ヌ） F 2014 07/15 01:46:44
2925 13 （セ） -> （ネ） F 2014 07/15 01:58:06
2933 13 （セ） -> （ノ） F 2014 07/15 02:04:15
9794 50 （セ） -> （ハ） S 2014 07/15 17:36:02
2767 13 （ソ） -> （ヒ） F 2014 07/16 08:54:32
1713 15 （ソ） -> （フ） F 2014 07/16 09:13:14
2767 13 （ソ） -> （ヒ） F 2014 07/16 09:19:58
2743 13 （ソ） -> （ヘ） F 2014 07/16 09:40:49
2767 13 （ソ） -> （ヒ） F 2014 07/16 09:44:59
122862 448 （ソ） -> （ホ） S 2014 07/16 14:14:21
```

(c) 侵入されたと考えられる SSH クライアントの通信ログ

図 7 重複検知した送信元のコネクション
Fig. 7 Connections from duplicate detected source IP addresses.

5. 提案手法の運用結果と考察

4 章で検証した結果より SCRAD-SS を 2015 年 2 月より本格的に運用開始した．そこで，2015 年 2 月から 2016 年 5 月までの SCRAD-SS の運用結果を分析し，提案手法の検知精度を評価した．

表 9 運用期間中の検知結果
Table 9 Detection results during operational periods.

(a) SSH クライアント検知数

クライアントの種類	攻撃者	正規ユーザ	重複検知	合計*
検知数	5,852	2,361	(114)	8,099

* 合計 = 攻撃者数 + 正規ユーザ数 − 重複検知数

(b) SSH コネクション判定数

通信の種類	判定数
SUCCESS コネクション	35,332
FAIL コネクション	270,513
合計	305,845

表 10 SCRAD-SS と真の値による分類におけるクライアント検知結果の関係
Table 10 Relationship between classification of true values and detection results of SCRAD-SS.

真の値＼検知結果	正規ユーザクライアント	攻撃者クライアント
正規ユーザ	α	β
攻撃者	γ	δ

5.1 検知結果

2015 年 2 月 1 日から 2016 年 5 月 31 日までの運用結果を**表 9** に示す．表 9 (a) にはクライアント検知結果を，表 9 (b) にコネクション判定数を示す．表 9 (a) に示すように，運用期間中に 114 件の送信元 IP アドレスを重複検知している．4.4.6 項で述べたように，重複検知は攻撃者としても正規ユーザとしても検知した送信元 IP アドレスである．このような送信元 IP アドレスは誤検知，検知漏れの疑いがある．重複検知した送信元 IP アドレスについては 5.3 節に調査結果を示す．表 9 (b) より，観測した全コネクション 305,845 件のうち，88.4%にあたる 270,513 件の FAIL コネクションを観測している．

5.2 検知精度の評価

検知精度の評価には 4.4.5 項で用いたコネクション判定率に加えて，SCRAD-SS による攻撃者の検知精度を評価するためクライアント検知率を算出する．クライアント検知率においても，コネクション判定率同様に SCRAD-SS による分類が真の値による分類と比較して，どれだけ一致しているかを攻撃者，正規ユーザにおいて算出する．**表 10** の α，β，γ，δ を用いて，以下の式より求める．

正規ユーザ検知率

$$\frac{\alpha}{\alpha+\beta}$$

攻撃者検知率

$$\frac{\delta}{\gamma+\delta}$$

表 11 運用結果における真の値
Table 11 True values during operational periods.

(a) SSH クライアント

クライアントの種類	真の値
真の正規ユーザ	1,947
真の攻撃者	6,978
不明なユーザ	7,704
合計	16,629

(b) SSH コネクション

通信の種類	真の値
正規ユーザによるコネクション	44,829
攻撃者によるコネクション	251,874
不明なユーザによるコネクション	9,142
合計	305,845

表 12 真の値による分類と検知結果の関係
Table 12 Relationship between classification of true values and detection results.

(a) クライアント検知結果と真の値の関係

真の値＼検知結果	正規ユーザ	攻撃者
真の正規ユーザ	1,947	414
真の攻撃者	19	5,833

(b) コネクション判定結果と真の値の関係

真の値＼検知結果	正規ユーザ	攻撃者
真の正規コネクション	32,352	12,477
真の攻撃コネクション	3	251,871

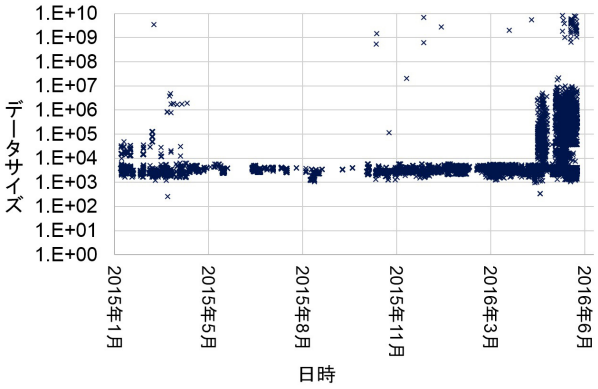
表 13 検知精度の調査結果
Table 13 Results of detection accuracy.

評価指標	コネクション (%)	クライアント (%)
攻撃者	99.9	99.7
正規ユーザ	72.2	82.5

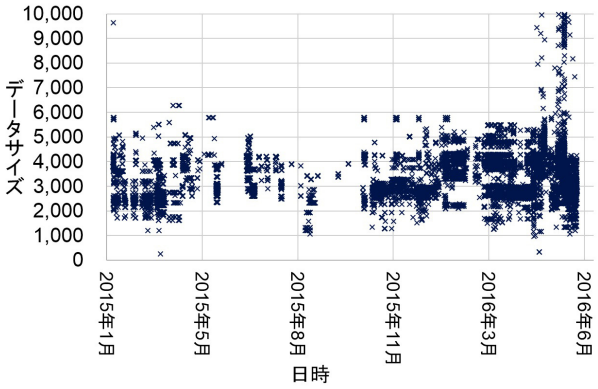
まずは運用期間中におけるすべての送信元 IP アドレスについて真の値を調査した結果を**表 11** に示す．ここで述べる真の値は 4.2.2 項で述べた真の値の算出方法により求めた．次に，真の値による分類と運用期間中における検知結果の関係について調査した結果を**表 12** に示す．なお，表 12 では不明なユーザを省いている．表 12 に基づいてコネクション判定率，クライアント検知率それぞれを算出した結果を**表 13** に示す．これらの結果から，攻撃者のコネクション判定率については 4.4 節の実験結果（表 8 SCRAD-SS の列項）よりも向上している一方，正規ユーザのコネクション判定率は減少したことが判明した．

5.3 重複検知した送信元によるコネクションの調査

運用期間中に観測したコネクションの中で，誤判定の可



(a) 重複検知した送信元のコネクションにおけるデータサイズ



(b) 重複検知した送信元のコネクションにおけるデータサイズ（0〜10,000 バイト）

図 8 重複検知した送信元によるコネクション
Fig. 8 Connections from duplicate detected source IP addresses.

```
5069 18 （マ） -> （ミ） 2015 04/18 06:13:37
5069 18 （マ） -> （ミ） 2015 04/18 06:13:56
5069 18 （マ） -> （ミ） 2015 04/18 06:14:04
5069 18 （マ） -> （ミ） 2015 04/18 06:14:16
5069 18 （マ） -> （ミ） 2015 04/18 06:14:26
5069 18 （マ） -> （ミ） 2015 04/18 06:14:34
5069 18 （マ） -> （ミ） 2015 04/18 06:14:41
5069 18 （マ） -> （ミ） 2015 04/18 06:14:47
5069 18 （マ） -> （ミ） 2015 04/18 06:15:10
5069 18 （マ） -> （ミ） 2015 04/18 06:15:32
```

(a) 攻撃者を検知漏れしたと考えられる通信ログ

```
2553 8 （ム） -> （メ） 2015 12/02 16:07:30
3753 19 （ム） -> （メ） 2015 12/02 16:09:11
3153 13 （ム） -> （メ） 2015 12/02 16:10:20
3237 14 （ム） -> （メ） 2015 12/02 16:10:58
3037 12 （ム） -> （メ） 2015 12/02 16:11:18
2553 8 （ム） -> （メ） 2015 12/02 17:24:23
2553 8 （ム） -> （メ） 2015 12/02 18:01:35
3637 18 （ム） -> （メ） 2015 12/02 18:01:51
2553 8 （ム） -> （メ） 2015 12/02 18:30:47
2553 8 （ム） -> （メ） 2015 12/02 18:31:53
20788097 24596 （ム） -> （メ） 2015 12/02 19:32:59
```

(b) 正規ユーザを誤検知したと考えられる通信ログ

図 9 重複検知した送信元のコネクション
Fig. 9 Connections from duplicate detected source IP addresses.

能性が高いものに関して本節で調査結果を述べる．重複検知した送信元 IP アドレス計 114 件によるコネクションを計 27,790 件観測した．**図 8** は重複検知した送信元のコネクションにおけるデータサイズと日時の関係を表しており，コネクションがまったく観測されていない部分は SCRAD-SS が運用を停止した期間である．図 8 (a) から，重複検知した送信元 IP アドレスによるコネクションは大量のデータをやりとりするコネクションをつないでいるものもあるが，27,790 件中 24,051 件が 1 コネクションに送受信するデータサイズが 1,000 と 10,000 の間に存在している（図 8 (b)）．このことから，主に観測した SUCCESS コネクションがコネクション判定のしきい値（5,000 バイト）をわずかに上回るコネクションのみであるかどうかで，正規ユーザの誤検知か攻撃者を検知漏れしたかを判断した．しかし，大量のデータを送受信したコネクションをつないだ送信元 IP アドレスについても，学内の SSH サーバに侵入された可能性がある．重複検知した送信元のうち 2 件のコネクションの一部を抜粋した（**図 9**）．図 9 (a) は送信元が攻撃者にもかかわらず検知漏れしたと判断した通信ログである．攻撃者と判断した根拠としては，図 9 (a) の左から 2 番目のフィールドが指す $PN_{detection}$ が 18 パケットであり（SCRAD-CP では攻撃者として検知），平均 12 秒とコネクションの接続間隔が短いため正規ユーザの挙動とは考えにくいためである．しかし，SSH を用いてファイルを送受信するコマンド（scp，rsync など）を用いた場合，キーインタラクションによるパケットのやりとりがないためパケット数が少なくなる可能性がある．また，定期的にコマンドやプログラムを実行する cron コマンドを用いた場合，コネクション接続間隔についても短くなる可能性も考えられる．以上のことから，図 9 (a) の送信元を攻撃者と断定することはできないため，scp や rsync など SSH を用いるコマンドによって生じるコネクションについては今後さらに調査する必要がある．

一方，図 9 (b) では大量のデータを送受信していることから，確実に SSH サーバとデータを送受信していると判断したため，正規ユーザを誤検知したと判断した．

5.4 運用結果の考察

運用結果を分析したことにより，大分大学における SSH パスワードクラッキング攻撃の 99.9%を検知し，その送信元 IP アドレスからの通信を遮断できたことが判明した．

一方で，重複検知した送信元について調査すると，正規ユーザの誤検知と考えられるコネクションを観測している．よって，SCRAD-SS により SSH パスワードクラッキング攻撃による SSH サーバへの侵入のリスクを低減できることと，正規ユーザを誤検知することにより，ユーザの利便性を損ねる可能性があることが判明した．

6. おわりに

6.1 まとめ

パケット数を用いた検知手法では，scp コマンドを用いたデータの送受信の際に，正規ユーザを攻撃者として誤検知するという問題点があった．

この誤検知を改善するため，データサイズを用いた検知手法を提案した．scp でデータの送受信をする場合の最小値と，攻撃者の通信のデータサイズの最大値から，データサイズを用いた検知手法における攻撃検知のしきい値として，5,000 バイト未満のコネクションを FAIL コネクションとした．

データサイズを用いた検知手法と，パケット数を用いた検知手法を，データセットを用いて比較した．送信元 IP アドレス別に検知精度を比較した場合は，誤検知を 1 件改善できた．また，攻撃者と正規ユーザによるコネクションそれぞれの検知率を調査したところ，どちらの検知手法でも，攻撃者によるコネクションの検知率は 99.7%であった．一方，正規ユーザによるコネクションの検知率は，パケット数を用いた検知手法では，67.3%であったが，データサイズを用いた検知手法では，90.1%であった．

2015 年 2 月から 2016 年 5 月まで SCRAD をデータサイズを用いる検知手法で運用した結果を分析したところ，攻撃者の検知精度が 99.9%検知できた一方で，正規ユーザの検知精度は 4.4 節の実験結果よりも低く，82.5%であった．SCRAD-SS の運用により，2015 年 2 月から 2016 年 5 月までの運用期間中に 5,833 件の攻撃者による 251,871 件のパスワードクラッキング攻撃による SSH サーバへの不正侵入を防ぐことができた．

6.2 今後の課題

正規ユーザによるコネクション判定率や，クライアント検知率を改善するために，SSH を用いるコマンド（scp,rsync など）によって生じるコネクションについて詳細に調査する必要がある．これまでは公開鍵認証を用いる前提で，scp や rsync コマンドにより 1 コネクションで送受信するデータサイズについて調査してきた．今後は，パスワード認証を用いる場合や，その他の SSH を用いるコマンドを利用した場合についても調査する必要がある．

本論文では，SSH コネクションにおいて送受信するデータサイズに基づく検知手法を提案し，運用・評価した．しかし，SSH プロトコルにおける鍵交換の段階ではパケットの大きさに特に制限はないため，この段階で極端に大きな鍵データを送るなど，SCRAD-SS の検知を回避する手段は存在する．このような回避手段への対策は今後の課題である．

参考文献

[1] JPCERT インターネット定点観測レポート（2015 年 1〜3 月），入手先 ⟨http://www.jpcert.or.jp/tsubame/report/report201501-03.html⟩.

[2] JPCERT インターネット定点観測レポート（2016 年 1〜3 月），入手先 ⟨https://www.jpcert.or.jp/tsubame/report/report201601-03.html⟩.

[3] インターネット観測システム（TSUBAME）観測グラフ，入手先 ⟨https://www.jpcert.or.jp/tsubame/graph.html⟩.

[4] 清水光司，小刀稱知哉，池部　実，吉田和幸：再送パケット除去による SSH パスワードクラッキング攻撃検知システムの検知方法の改善，第 67 回電気・情報関係学会九州支部連合大会，p.87 (2014).

[5] 小刀稱知哉，中本菜桜美，清水光司，池部　実，吉田和幸：SSH パスワードクラッキング攻撃検知システムの改善とその運用結果, 情報処理学会研究報告（インターネットと運用技術），Vol.2014-IOT-26, No.4, pp.1–7 (2014).

[6] Vykopal, J., Plesnik, T. and Minarik, P.: Network-Based Dictionary Attack Detection, *2009 International conference Future Networks*, pp.23–27 (Mar. 2009).

[7] Nagios, available from ⟨http://www.nagios.org/⟩.

[8] 大隅淑弘，山井成良，井上一郎二：アクセス制御ファイルの動的変更による SSH 総当り攻撃への対策，学術情報処理研究，No.11, pp.68–73 (2007).

[9] 小刀稱知哉，天本大地，池部　実，吉田和幸：SSH パスワードクラッキング検知システムその遮断の効果について，情報処理学会マルチメディア，分散，協調とモバイル（DICOMO2013）シンポジウム，pp.742–748 (2013).

[10] 清水光司，小刀稱知哉，池部　実，吉田和幸：ボットネットによる SSH パスワードクラッキング攻撃の検知のための予備調査，情報処理学会研究報告（インターネットと運用技術），Vol.2014-IOT-29 (2015).

清水 光司　（学生会員）

平成 27 年大分大学工学部知能情報システム工学科卒業．同年同大学大学院工学研究科知能情報システム工学専攻博士前期課程に進学し，現在，在学中．ネットワークセキュリティに興味を持つ．

小刀稱 知哉

平成 25 年大分大学工学部知能情報システム工学科卒業．平成 27 年同大学大学院工学研究科知能情報システム工学専攻博士前期課程修了．現在，三菱電機株式会社に勤務．

池部 実 （正会員）

平成 16 年大分大学教育福祉科学部情報社会文化課程卒業．平成 18 年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了．平成 23 年同大学情報科学研究科博士後期課程修了．同年筑波技術大学保健科学部特任助教を経て，平成 24 年大分大学工学部知能情報システム工学科助教，現在に至る．博士（工学）．ネットワーク運用技術，ネットワークセキュリティ，広域分散処理システムの研究に従事．電子情報通信学会，ACM，IEEE 各会員．

吉田 和幸 （正会員）

昭和 54 年九州大学工学部情報工学科卒業．昭和 59 年同大学大学院工学研究科情報工学専攻博士後期課程修了．同年大分大学工学部講師，昭和 61 年同助教授，平成 14 年同総合情報処理センター助教授を経て，平成 20 年同学術情報拠点教授．工学博士．ネットワークの運用技術，セキュリティに関する研究に従事．電子情報通信学会，ソフトウェア科学会，ACM，IEEE 各会員．