Vol.2018-CSEC-81 No.17
Vol.2018-IOT-41 No.17
2018/5/18

Darknetの解析に基づくSSH攻撃傾向の分析

小林 孝史[1,a)]　俣野 剛志[1,†1]　坂東 翼[2,†2]

概要：インターネットからアクセス可能な機器が増加傾向にあり，攻撃・被害に遭う可能性が高くなっている．管理の行き届いていないサーバや，デフォルトのIDとパスワードで運用している機器等が特に狙われると見られており，その傾向をできるだけ早期に把握し，サーバの運用方法を変えるなどの対策が必要である．
当研究室ではSSHハニーポットを運用しており，パスワードの総当り攻撃等を防止する研究に取り組んでいる．より詳細な攻撃傾向を掴むために，Darknetへのアクセスログを解析し，DarknetとSSHハニーポットへのアクセスの相関関係等を分析した．その結果，Darknetへのアクセス後約1日以内にSSHハニーポットへのアクセスが発生し，そのほとんどが14時間以内に集中していることが分かった．

Analysis of SSH attack tendency based on Darknet analysis

Takashi Kobayashi[1,a)]　Tsuyoshi Matano[1,†1]　Tsubasa Bando[2,†2]

Abstract: The number of devices that can be accessed from the Internet is on the rise, and the possibility of being attacked / damaged is increasing. It seems that servers that are not well managed and devices operating with the default ID and password are specifically targeted, and the administrator can grasp the trend as soon as possible and change the operation method of the server Measures are necessary.
Our laboratory operates SSH honeypot, and we are working on research to prevent password brute force attack etc. In order to grasp a more detailed attack tendency, we analyzed the access log to Darknet and analyzed the correlation between access to Darknet and SSH honeypot. As a result, access to the SSH honeypot occurred within about 1 day after accessing Darknet, and most of it was found to be concentrated within 14 hours.

1. はじめに

インターネットは，個人が企業が当たり前に日常的に利用し，社会生活に不可欠のものとなっている．しかし，悪意ある利用者による攻撃が後を絶たず，近年では攻撃の精度は高度化し，攻撃の種類も巧妙化している．また，特定の機関や組織を狙った攻撃のみならず，すべてのユーザに対して無作為に攻撃が行われている [1]．

最近では，IoTデバイスが多く存在し，シェルアカウントを持ったIoTデバイスが不正アクセスの目標として狙われることが多くなっている．標準のIDとパスワードの組み合わせのまま運用しているデバイスも多く存在すると予想されており，設定を書き換えられたり，他のサイト等を攻撃する踏み台にされたりする事例も多くなっている．

シェルアカウントが有効なUNIX系のOSやIoTデバイスでは，TELNETやSSHプロトコルが用いられており，パスワードクラック攻撃等の対象となりやすい．したがって，これらの環境におけるID・パスワード管理は確実に行なう必要はあるが，実際には弱いパスワードで運用を行っている箇所も少なくない．

こういったID・パスワード管理の甘い環境は，パスワードクラック攻撃には非常に弱いものとなる．そのような環境を攻撃から守るためには，その機器自体の管理を確実に行うことも必要であるが，ネットワーク側でもその傾向を掴んだ上で攻撃元からの接続の試みを遮断する必要も出て

[1] 関西大学 総合情報学部
Faculty of Informatics, Kansai University
[2] 関西大学大学院 総合情報学研究科
Graduate School of Informatics, Kansai University
[†1] 現在，NECソリューションイノベータ株式会社
[†2] 現在，株式会社インフォセック
[a)] taka-k@kansai-u.ac.jp

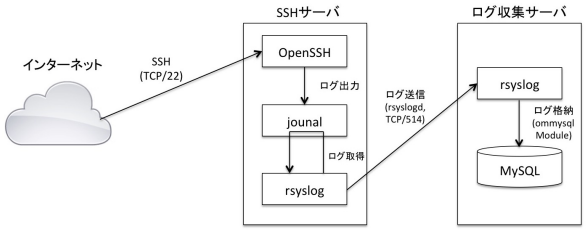

図 2 SSH ログ取得のシステム構成図

くると考えている．

そのために，攻撃元からのアクセスの傾向を掴み，それを実際に運用している機器へのアクセスの試みにどの程度影響しているのかを検証し，対策の指針としたい．

2. 関連研究

深澤らの研究 [2] では，日本の Darknet を観測している NICTER [3] と世界規模の Darknet を集めている NORSE の二つの Darknet 観測網を用いてトラフィックデータの相関分析を行い，両方の Darknet トラフィックに相関関係があることを示した．さらに，Darknet トラフィックを分析することにより，ターゲットとされているポートや地域を推測することが可能であることを示した．この研究では相関関係を調査するために，日別での検知数，特定日の時間別での検知数，IP アドレス別での検知数，世界規模でのニュースとなった事件との相関分析の四つの観点から調査を行った．

本研究では，深澤らの IP アドレス別での検知数を用いた検証を参考に Darknet と Livenet に両方にアクセスした IP アドレスを算出し，それぞれのアクセスの日時を比較検討を行う．

坂東・上原らの研究 [4] では，認証時間に基づいたパスワードクラッキング攻撃検知機能を既存の OpenSSH サーバに実装し，実際のパスワードクラッキング攻撃に対する提案手法の有効性を検証した．また，小林研究室内に SSH ハニーポットを設置し，アクセスをデータベースでまとめ，関西大学宛に行われた SSH パスワードクラッキング攻撃を分析した．そのデータベースへの登録内容は図 1 のようなものである．

図 2 に坂東・上原らの研究システムの構成図を示す．図 2 に示される SSH サーバは，22/TCP へのアクセスを受け付け，送信元 IP アドレスやアクセスが行われた日時に加え，認証に要する時間や攻撃判定結果を含むアクセス情報をログとして出力する．出力されたアクセス情報のログは rsyslogd を用いてログ格納サーバに格納される．そして，ログ格納サーバ内の rsyslogd によって MariaDB に MySQL Database Output Module を用いて格納される．

本研究では，この SSH ハニーポット（以下，Livenet と呼ぶ）に対するアクセスを記録したデータベース内にまとめられた送信元 IP アドレスとアクセスが行われた日時を用いて，NICTER が保有するサイバーセキュリティ情報の分析基盤 NONSTOP システム [5] を介して得られた Darknet へのアクセス時間との差を比較検証する．

表 1 宛先番号別パケット観測数トップ 5

順位	宛先ポート番号	前四半期の順位
1	23/TCP(telnet)	1
2	22/TCP(ssh)	3
3	445/TCP(microsoft-ds)	4
4	1433/TCP(ms-sql-s)	2
5	2323/TCP	6

3. 近年の攻撃先ポート番号の推移

JPCERT/CC が発表している 2017 年 10 月 1 日から 12 月 31 日までの「インターネット定点観測レポート [6]」では，宛先ポート番号別のパケット観測数のトップ 5 を示している．表 1 に宛先番号別のパケット観測数の上位五つをまとめたものを示す．本研究で対象としている SSH（22/TCP）は 2 位となっており，定点観測でもアクセスが多いことが分かる．

`telnet` や `ssh` へのアクセスが観測数トップを占めており，依然として攻撃の対象となりやすいことがわかる．攻撃者がシェルアカウントを攻撃の対象として，そのシステム等へのパスワードクラック攻撃ののちに侵入を試み，パスワード管理の甘いユーザで侵入後活動を行なう．その侵入の試みを事前に検出できれば，早期警戒情報としてブラックリストを提供し，そのリストに基づいて不正アクセスを防止できる可能性が出てくる．

本研究の目的は，そのブラックリストを作成する元になる Darknet と Livenet のアクセス時間の差を求めることにある．

4. Darknet のログと SSH ハニーポットのログの分析

4.1 アクセスログの分析の目的

アクセスログの分析は，Darknet にアクセスを試みたのちに，Livenet に対してアクセスを試みている IP アドレスがどのくらいあり，アクセス日時にどのくらいの差があるかを確認するために行う．

Darknet と Livenet の両方にアクセスしている一 IP アドレスが一番はじめにアクセスした日時を比較し，Darknet にアクセスした後に Livenet にアクセスした場合，Darknet にアクセスした IP アドレスを防ぐことによって不正アクセスのタイミングを予想できると考えている．

予備調査として，Darknet へのアクセスと Livenet へのアクセスの前後関係を調査した．赤丸を Darknet にアクセスした後に Livenet にアクセスした IP アドレスとし，青丸を Livenet にアクセスした後に Darknet アクセスした IP

Auth	User	IP	Time	Detect	RTT	Country	Year	Month	Day	Hour	Minute	Second	MicroSec	Datetime
Fail	service	197.211.[redacted]	10.415352	Normal	0.858614	ZW	2017	9	29	0	11	38	857542	2017-09-29 00:11:38.857542
Fail	service	197.211.[redacted]	0.405727	Attack	0.858614	ZW	2017	9	29	0	11	39	263457	2017-09-29 00:11:39.263457
Fail	avis	197.221.[redacted]	10.12497	Normal	0.661421	ZW	2017	11	2	16	46	14	319765	2017-11-02 16:46:14.319765
Fail	ubuntu	197.221.[redacted]	0.845068	Normal	0.895321	ZW	2017	11	2	20	31	31	217490	2017-11-02 20:31:31.217490
Fail	bot	41.78.[redacted]	4.961822	Normal	1.171823	ZW	2017	11	10	20	3	6	314261	2017-11-10 20:03:06.314261
Fail	duci	41.78.[redacted]	0.457357	Attack	0.454662	ZW	2017	11	10	20	3	15	102143	2017-11-10 20:03:15.102143

図 1　データベースの格納例

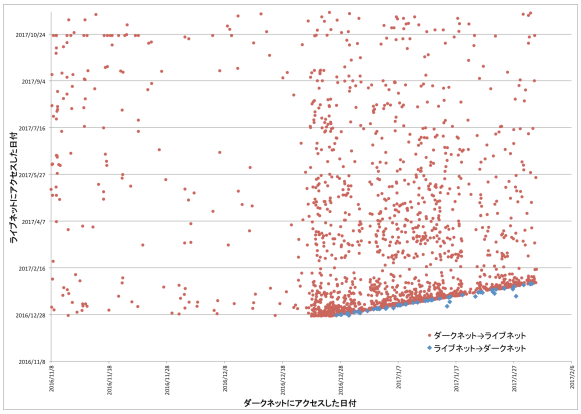

図 3　相互にアクセスした IP アドレスのアクセスする順番の分類：赤：Darknet → Livenet，青：Livenet → Darknet

アドレスとしたものを図 3 に示す．なお，Darknet へのアクセスは 2017 年 1 月 30 日以前のものである．

図 3 には赤丸がプロットされていることから，ほとんどの IP アドレスが Darknet にアクセスした後に Livenet にアクセスしたことが分かる．2016 年 12 月中旬以前の期間については，Darknet，Livenet ともにアクセスが非常に少ない状況であるが，それ以降についてはどちらにもアクセスするケースが多くなってきている．

特に，右下の赤丸が密集した部分は，Darknet へのアクセス日時と Livenet へのアクセス日時が非常に近い IP アドレス群が集まっており，比較的短時間で Darknet から Livenet へのアクセスへ遷移していることが分かる．

4.2 Livenet および Darknet にアクセスした IP アドレスの分析

Livenet のデータ取得開始日である 2016 年 11 月 8 日から 2017 年 11 月末日の期間内で，同一 IP アドレスがそれぞれの空間に最初にアクセスした日時を比較した．

まず，Livenet に対して，一度でもアクセスを試みた IP アドレスを確認する．また，Darknet のポート番号 22 に対して，Livenet のデータの記録が開始された 2016 年 11 月 8 日から 2017 年 11 月末日までに一度でもアクセスした IP アドレスを確認する．

確認した値は，Darknet_IP_list に保存する．そして，Livenet 内で確認された値（Livenet_IP_list）が Darknet でも出現するかどうかを確認する．

確認には Python のモジュールである Pandas[7] を用いる．Pandas はデータ解析を支援する機能を提供するモジュールであり，時系列データなどを操作するために用いられる．

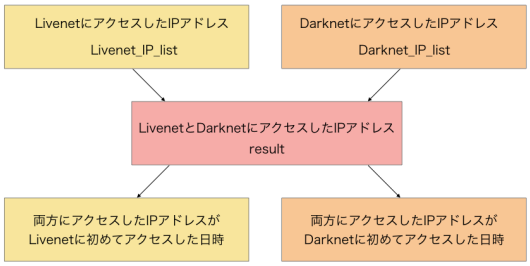

図 4　解析方法のフローチャート

そして，Darknet と Livenet に共にアクセスした IP アドレスの一覧を result に保存する．result.txt を用いて，両方にアクセスした IP アドレスが Darknet と Livenet にはじめにアクセスした日時を確認した．確認には Python のモジュールである MySQLdb[8] を用いる．MySQLdb は，Python で MySQL を使用するために用いる．MySQLdb.connect() でデータベースに接続を行い，引数には，ユーザー名，パスワード，ホスト名，ポート番号，接続するデータベースを渡す．MySQLdb のモジュールはインストールする必要がある．また，フローチャートにしたものを図 4 に示す．

5. 結果と考察

過去に Livenet にアクセスしたことのある IP アドレスは 14151 件であり，Livenet と Darknet の両方でアクセスが見られた IP アドレスは 2294 件であった．2294 件のうち，194 件が Livenet 内でアクセス日時を取れていなかったため，研究データとして扱わず，2100 件の IP アドレスを研究データとして扱う．Darknet にアクセスした後に Livenet にアクセスした IP アドレスは 2041 件あった．

それぞれにアクセスを試みた IP アドレスと，その IP アドレスが初めてアクセスを試みた時間を算出したものの一部を表 2 に示す．Livenet のアクセスログに記録されている日時，Darknet のログに記録されている日時ともに秒までが記録されており，秒単位での差分を取ることが可能である．

Darknet にアクセスした後に Livenet にアクセスした IP アドレスについて，それぞれの空間へのアクセス日時の差がどれくらいになっているかを集計したグラフを図 5 に示す．図 5 より，日数が増えるごとにその度数が減っている

表 2　Darknet と Livenet 両方にアクセスがあった IP アドレスとそれぞれのアクセスを試みた日時

IP アドレス	Darknet	Livenet
22x.xxx.xxx.xxx	2016/12/23 0:42	2016/12/26 12:01
11x.xxx.xxx.xxx	2016/12/23 10:58	2016/12/26 12:12
6x.xxx.xxx.xxx	2016/12/26 4:24	2016/12/26 12:13

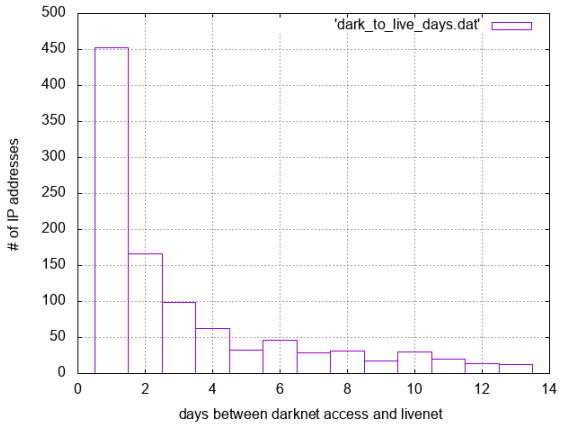

図 5　Darknet へアクセス後，Livenet へアクセスするまでの日数

ことがわかる．このことは，図 3 で示した予備実験の結果を反映している．つまり，Darknet へアクセスが発生したあと，短時間で Livenet へのアクセスが多く発生しており，長時間経過してから Livenet へアクセスすることは少なくなっていることを表している．

Darknet にアクセスした後に，1 日以内に Livenet にアクセスした IP アドレスが 453 件，2 日以内が 166 件，3 日以内が 99 件であった．このことから，多くの IP アドレスにおいて，Darknet にアクセスした後に Livenet にアクセスしていることが分かった．

次に，Darknet と Livenet に 1 日以内にアクセスした IP アドレスが何時間以内にアクセスしたものかを図 6 に示す．この図より，Darknet にアクセスした後，14 時間以内に Livenet にアクセスしている IP アドレスが多い傾向が見られる．また，この図の時間の範囲外ではあるが，24 時間以降も毎時間 10 件ほどの IP アドレスがアクセスしていることが観測されている．

6. おわりに

本論文では，SSH ハニーポット（Livenet）へのアクセスログと Darknet へ到達するアクセス試行のログを分析することで，Livenet への不正アクセスが到達する日数および時間についての傾向分析を行なった．

その結果，Darknet へアクセスした IP アドレスの 97%が Livenet へアクセスしており，そのうちの 21.5%が 1 日以内にアクセスを試みていることがわかった．さらに 14 時間以内に Livenet に到達していることが多く，Darknet へのアクセス試行が観測されてから 14 時間以内は Livenet での警戒が必要であるといえる．

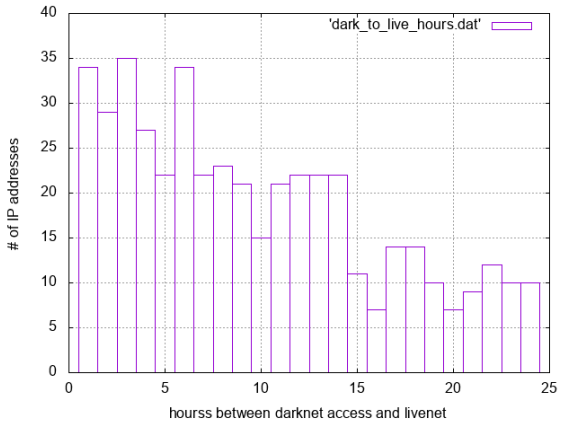

図 6　Darknet へアクセス後，Livenet へアクセスのあるまでの時間

SSH ハニーポットへのアクセスはほぼパスワードクラック攻撃のものである．パスワード認証時の時間的な解析はできており，不正なパスワード入力かどうかの判別はほぼできているが，不正と思われる接続自体の抑制にはつながっていない．本論文で述べた傾向分析によって，こういった不正な接続自体を抑制もしくは排除ができる可能性を示すことができたと考えている．今後，攻撃者側もこういった経過時間についての変化は加えてくると考えられるので，新たな対策を示していく必要があると考えている．

参考文献

[1] 警視庁，総務省，経済産業省：不正アクセス行為の発生状況（平成 28 年），http://www.soumu.go.jp/main_content/000473526.pdf, 2018 年 4 月 10 日確認．
[2] 深澤成孝, 佐藤直, “ダークネットトラフィックの相関分析”, 情報処理学会研究報告　IPSJ SIG Technical Report, Vol.2015-CSEC-68 No.20, 2015.
[3] 独立行政法人情報通信研究機構, “NICTER Darknet 2014”, http://www.iwsec.org/mws/2014/files/NICTER_Darknet_Dataset_2014.pdf, 2018 年 4 月 10 日確認.
[4] 坂東, 上原, 小林, “認証時間に基づいた SSH パスワードクラッキング攻撃検知手法の提案”, 第 16 回情報科学技術フォーラム, L-015, 2017.
[5] 竹久, 神薗, 笠間, 中里, 衛藤, 井上, 中尾, “サイバーセキュリティ情報遠隔分析基盤 NONSTOP の利活用について”, コンピュータセキュリティシンポジウム 2014 論文集, Vol.2, pp.207-214, 2014.
[6] JPCERT/CC ：インターネット定点観測レポート（2017 年 10〜12 月）, https://www.jpcert.or.jp/tsubame/report/report201710-12.html, 2018 年 4 月 10 日確認.
[7] Python Data Analysis Library pandas, http://pandas.pydata.org/, 2018 年 4 月 10 日確認.
[8] mysqlclient 1.3.12 : Python Package Index, https://pypi.python.org/pypi/mysqlclient/, 2018 年 4 月 10 日確認.