ファイル共有ソフトウェアの検出法の更新

上埜 元嗣

情報社会基盤研究センター

概要

本学ではセキュリティポリシーによりファイル共有ソフトウェアの利用を禁止しており、以前より使用者を検出してきた。技術の進歩に伴い検出方法も変わってきており、本年度、機器の更新に伴い検出方法が変わり、現在は運用しながら精度や効率を高めている。本稿ではそれに伴う問題点や課題について述べる。

1　はじめに

2002 年より以前から試行していたファイル共有ソフトウェアの検出を正式に始めた。また、2003 年にはセキュリティポリシーも策定され、その中でもファイル共有ソフトウェアの使用禁止に関する条項はもりこまれた。ファイル共有ソフトウエアの使用を禁止するためにはファイヤーウォールでの通信を遮断することで使用不可とすることも可能であったが行わなかった。当時は技術的にファイル共有ソフトウェアの通信のみを遮断することは難しく、そのほかの通信に支障をきたすことも十分あったうえ、本学は研究機関であるため研究目的での使用もあるためである。そのため、通信の検出からユーザを特定し、ユーザに注意喚起という手段をとっている。検出機器も更新しながら行っており、今年度はちょうど更新した。本稿ではいくつかの機器更新のうち FortiNet 社 FortiGate3950B の更新に伴う問題点や課題について更新作業中であるが報告する。

2　機器の更新

2002 年より始めた検出では tcpdump によりネットワークのパケットをキャプチャしその中でファイル共有ソフトウェアの通信に使われる port を使用したパケットをカウントしカウントの多い IPaddress を利用者と断定していた。しかしながら解析しなければならないデータが膨大なことや解析に時間がかかるために即座に検出することができないことが問題であった。その後の機器の更新では Lancope 社 Stelth Watch System を導入しネットワークのトラフィック管理によって検出を行った。このシステムでは netflow,sflow によりトラフィックデータを収集し分析までおこなえ、グラフィカルにそのデータを閲覧できた。検出のための作業効率はよくなった。今回の更新ではファイヤーウォールとして FortiNet 社 FortiGate3950B 、トラフィック管理として Genie 社 Genie6333-T、パケットキャプチャとして FLUKEnetworks 社、NetworkTimeMachine Express3 を導入した。それぞれはファイル共有ソフトウェアの通信の検出が目的ではなくそれぞれに主の目的があり機能としてファイル共有ソフトウェアの通信の検出が可能である。

2.1　ForiGate3950B

FortiNet 社 FortiGate3950B はファイヤーウォールの機器更新に伴い導入された。主な目的はもちろんファイヤーウォールであり機能も当然ファイヤーウォールとしての機能が充実している。主な特徴としては

- 高性能ハードウェア　　最大 20Gbps（本学仕様）のパフォーマンス
- モジュラー型の拡張性
- 複合脅威セキュリティー

2.2　複合脅威セキュリティー

従来はファイヤーウォールとは別にアンチウィルス、アンチスパム、web フィルタリング、IPS などのシステムを導入しファイヤーウォールと連携させることで制御してきた。FortiGate はファイヤーウォールであるが、それらの機能を持っており、それによりコスト削減、負荷軽減、耐障害性の向上などをうたっている。

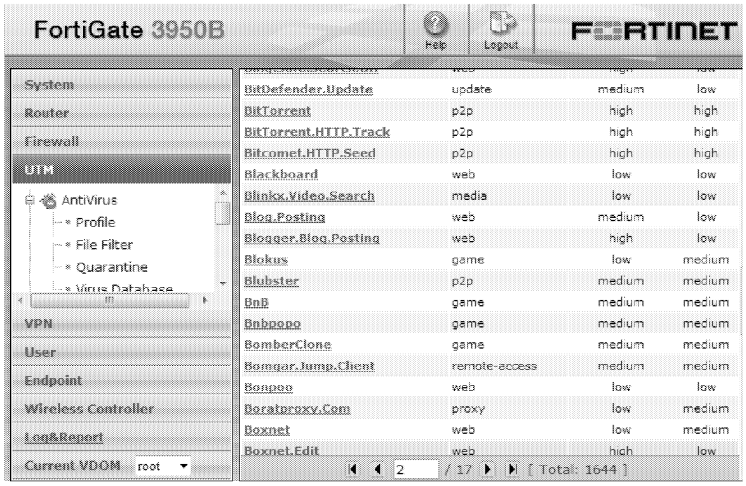

図1.　FortiGate3950B　Application list

これらの機能の一部として 1400 以上のアプリケーションの通信を認識し制御できる。もちろん、Winny や Share, BitTorrent など国内外のファイル共有ソフトウェアの通信を認識できる。ファイル共有ソフトウェアについては約 80 のソフトウェアを認識できる。

3　設定およびソフトウェアの検出

今回は更新作業の初段階ということから FortiGate3950B がどれくらい正確にファイル共有ソフトウェアを検出できるかに重点を置いて設定した。

3.1　FortiGate3950B の設定

以下のポイントを踏まえ設定した。

- 登録されているファイル共有ソフトウェアはすべて検出する
- ファイル共有ソフトウェアの通信は遮断しない
- 通信の検出はログをとる

図 2 は FortiGate3950B に対し上記の条件で設定を行っているところである。

図2.　FortiGate3950B　Appliation Control の設定画面

3.2　ログ

ログは FortiGate 本体に記録しておくこともできるがサイズの制限があるためログサーバに転送することにした。本学のシステム全般のログを収集しているログサーバに転送する設定をした。図 3 にて実際のログを示す。

ログからはファイル共有ソフトウェアを使用している IPaddress や相手先の IPaddress 使用ポート使用ソフトウェアなどが記述されている。このようなログが 1 日あたり 400 万行程度あり、それらのを 1 行づつ分析するわけにはいかないので、集計するプログラムを作成した。


```
2011-07-11 01:00:00 150.65.254.129 (20.6) <166>date=2011-07-11,time=00:59:59,devname=fg39a
,device_id=FG3K9B3E10700172,log_id=1059028704,type=app-ctrl,subtype=app-ctrl-all,pri=infor
mation,vd="root",attack_id=0,user="N/A",group="N/A",src=150.65.246.70,src_port=11763,src_i
nt="vlan3016",dst=90.214.135.56,dst_port=56881,dst_int="vlan3001",src_name="150.65.246.70"
,dst_name="90.214.135.56",profilegroup="N/A",profiletype="N/A",profile="N/A",proto=17,serv
ice="56881/udp",policyid=11951,serial=1945523475,app_list="p2p",app_type="p2p",app="BitTor
rent",action=pass,count=1,msg="N/A"
```


図3.　FortiGate3950B のログ

3.3　集計プログラム

ログデータの中で今回必要なものを挙げる

- 学内で使用している IPaddress(sorce,distnation のどちらの場合も)
- 使用しているファイル共有ソフトウェア
- 検出したログ数
- 接続先の IPaddress 数

今回は、検出の正確さに重点を置いているので、対象がどのようなふるまいをしているかということも重要と考えて検出の確かさを判断する。また、一つ一つのソフトウェアの挙動や複数使用しているかどうかなどの判断のためにも学内で使用している IPaddress とファイル共有ソフトウェアをキーにしてそれぞれの検出数を出した。プログラム言語は perl を使用し、ログサーバ上で実行した。

FortiGate P2P Summary Report

IP	P2P_App	Log Count	Distnation_IP_number
150.65.1.130	Sina.TV	3	1
150.65.1.131	Sina.TV	5	1
150.65.1.1	Sina.TV	233	4
150.65.102.103	BitTorrent	267	2
150.65.104.106	BitTorrent	170105	27383
150.65.104.106	BitTorrent.HTTP.Track	881	13
150.65.104.106	Gnutella	7	7
150.65.104.112	Skype	8	5
150.65.104.118	Skype	5	3
150.65.104.57	PPLive	2	2
150.65.105.103	BitTorrent	11	1
150.65.105.108	Thunder	22	12
150.65.105.109	Skype	1	1
150.65.105.118	BitTorrent	28	1
150.65.105.17	Skype	17	13
150.65.108.103	Skype	30	11
150.65.108.106	BitTorrent	2	2
150.65.108.108	Skype	802	15
150.65.108.114	Skype	6	5
150.65.108.120	Skype	1	1
150.65.108.124	Skype	376	14
150.65.108.129	Skype	85	22
150.65.108.152	Skype	6	3
150.65.108.154	Skype	7	5
150.65.108.155	Skype	7	2
150.65.108.167	Skype	21	7
150.65.108.18	BitTorrent	9	1
150.65.108.16	Skype	1473	14
150.65.108.172	Skype	22	10

図4. 集計プログラムの出力例

4　分析

集計結果をファイル共有ソフトウェアの使用割合でグラフにしたものを図 5 に示す。ただし Skype を除いた検出数の多いものをグラフにした。eDonkey、BitTorrent がほとんどの割合を占めることがわかる。また、集計結果の IPaddress から本学 DNS サーバなどファイル共有ソフトウェアをしようしていないはずの IPaddress からファイル共有ソフトウェアの通信を検出していることが分かった。DNS サーバからの検出、BitTorrent、eDonkey についてさらに分析をすすめてみた。

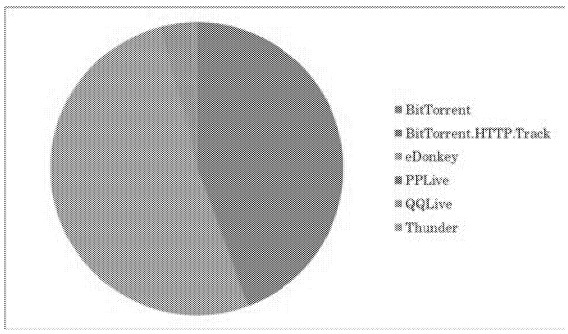

図5. ファイル共有ソフトウェアの割合

4.1　DNS サーバからの検出

本学 DNS サーバから Sina.TV というファイル共有ソフトウェアの通信を検出した。図 4 でもわかるが、最初の 3 つの IPaddress(150.65.1.130,150.65.1.131,150.65.1.1)は本学の DNS キャッシュサーバである。実際にファイル共有ソフトウェアを使用している可能性はない。そこで実際のログを確認したところあるきまった IPaddress の 53 番ポートに対してのアクセスであった。やはりそのことからこの検出については DNS のキャッシュサーバとしての通信をそのように誤検出してしまったものと考えられる。

4.2　BitTorrent

BitTorrent を検出した IPaddress では多数のログまた多数の通信先を検出していた。これはファイル共有ソフトウェア通信の特徴である。また、BitTorrent が検出された IPaddress からは BitTorrent 以外のソフトウェアも検出されている。それらのソフトウェアは BitTorrent を使用するものがほとんどである。IPaddress から使用者を特定し使用しているソフトウェアなどを確認してみると BitTorrent を使用しているとは思ってはいないが、そのほかの検出されたファイル共有ソフトウェアを使用していることを認めている。

4.3 eDonkey

eDonkey も検出割合では BitTorrent と同程度検出されている。しかしながら、特定の IPaddress からの検出が多いわけではなく数多くの IPaddress からの検出が多くまた、それぞれの IPaddress 検出数は少ない。これと検出数の分布が似ているものに Skype があり、IPaddress から使用者に確認を取ると Skype のみの使用ということであった。また、検出時刻なども Skype を使用していた時刻ではないが、常時 Skype を起動しているとのことであった。このことから Skype の何らかの通信を eDonkey と誤検出することが分かった。

5 考察

今回はファイル共有ソフトウェアの通信検出において FortiGate3950B の正確さを検証したわけであるが、以下のことが問題点であることが分かった。

- DNS サーバの特定サイトへの通信を誤検出してしまう。
- eDonkey としての検出はほぼ Skype の何らかの通信を誤検出してしまう。

特定の IPaddress に対する問題点は FortiGate3950B の設定で除外することは可能であり、設定したい。また、eDonkey の問題は Skype の挙動などを調査し何をどう誤認しているかを特定していきたい。統計データの方からも関連性を詳しく長期にわたりだし、Skype との関連が確かであれば Skype と同様の扱いとしたい。

使用率の少ないソフトウェアに関しても検証が必要かと思われる。

6 まとめ

まだ更新作業中であり FortiGate3950B の検出に関する精度を上げていかねばならない。そのほかの機器も同様に活用し、合わせた検出法でさらに精度も上げていきたい。また、当然利用者にはファイル共有ソフトウェアの使用禁止を啓蒙していく必要もあるわけであるが、IPaddress からの利用者の割り出しにコストがかかっている点も改善したい。今回の作業でこれらの課題がわかり、これらの課題に対して作業を進めていきたい。