双方向通信に着目した暗号化 P2P トラヒックの解析手法の検討

鈴木将史[1], 梅村恭司[1], 阿部洋丈[1], 岡部正幸[2]

[1]豊橋技術科学大学情報工学系　[2]豊橋技術科学大学情報メディア基盤センター

1. はじめに

近年, Winny などの P2P ファイル共有アプリケーションでは情報流出や著作権侵害などの問題が発生している. また, このほかの問題として P2P ファイル共有アプリケーションによるネットワークトラヒックは大幅な増大を見せており, インフラへの多大な負担となっている[5]. このため, P2P ファイル共有アプリケーションのトラヒックを特定し, 上記の問題を抑止することが期待されている.

これをうけて P2P ファイル共有アプリケーションを特定するためにいくつかの研究がなされている. これら従来の解析手法ではパケットのペイロードやヘッダ情報を解析することで P2P トラヒックを弁別していた[3][4]. しかし, これらの解析手法では通信を暗号化されてしまった場合, P2P トラヒックを弁別することができないといった問題が生じる.

そこで本稿では, P2P クライアントの双方向通信に着目した P2P トラヒック解析手法[2]を参考に, 通信が暗号化された状態における P2P トラヒックの特定手法[1]を検討する.

2. P2P トラヒック弁別手法

2.1 双方向通信に着目した P2P トラヒック弁別手法

通常のサーバとクライアント間での通信は, クライアント側からサーバ側へ通信路の確立(3 ウェイハンドシェイク)を行う. しかし, 一部のピュア P2P の通信では双方向から 3 ウェイハンドシェイクを行い, 各ノード間でネットワークを構築する必要がある. また P2P 通信では各ノードが自由に参加, 離脱できるため, 頻繁にノード間で通信路を確立する必要がある. このように頻繁に双方向に 3 ウェイハンドシェイクを行う通信はサーバクライアント通信では存在せず、ピュア P2P での通信以外にはあまり存在しない. そのため, この双方向通信に着目することで P2P トラヒックを弁別することができる. しかし, 文献[2]の手法では TCP のヘッダ情報を用いて 3 ウェイハンドシェイクを判定しているため, 通信が暗号化されてしまった場合には P2P トラヒックを弁別することができないという問題がある.

2.2 暗号通信に対する P2P トラヒック弁別手法

暗号化された通信では, TCP のヘッダ情報がわからないため, 本稿では事前に調査した SYN パケットと SYN/ACK パケットの時間間隔 α と, 相互に 3 ウェイハンドシェイクが行われる時間間隔 β をもとに P2P トラヒックを弁別する. これらの時間間隔を図 1 に示す.

以上の特徴量を用いて P2P トラヒックを弁別する. まず 3 ウェイハンドシェイクの判定条件は SYN パケットと SYN/ACK パケットのパケット長が共に 48byte であることから, 48byte のパケットが送信されてから時間間隔 α の間に 48byte のパケットが返信された場合に 3 ウェイハンドシェイクであると判定する. その後, 時間間隔 β 以内に反対側のホストから再度上記の判定方法で 3 ウェイハンドシェイクを行っていることが判定されれば, P2P 通信であると判定する.

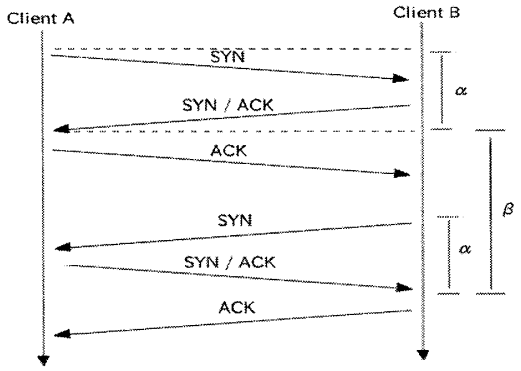

図 1：P2P トラヒックの特徴量

3. 評価実験

3.1 実験の流れ

本稿ではよく知られたピュア P2P ファイル共有アプリケーションである Winny を用いて実験を行う. 実験には暗号化されていない 2 種類のトラヒックを用いる. 1 つ目は Winny パケットのみを含むトラヒック(P2P トラヒック)であり, 2 つ目は Winny パケットを含まないトラヒック(nonP2P トラヒック)である. これらのトラヒックに対して弁別手法を適用し, 弁別手法の有効性を検討する. 実験の流れを以下に示す.

1. P2P トラヒックを後述する実験環境から取得し, nonP2P トラヒックを研究室のトラヒックから取得する. 各トラヒックの測定時間はそれぞれ 3 時間とした
2. 取得した P2P トラヒックから特徴量 α, β を決定
3. 決定した α, β を用いて P2P トラヒックの判定条件を設定し, P2P トラヒックと nonP2P トラヒックに対して上記の条件を適用した弁別手法を行い, 弁別手法の評価する

3.2 P2P トラヒック収集のための実験環境

実験環境にはホスト OS として Linux(Ubuntu 9.04)をインストールした 5 台のマシンを用意した. このうち 4 台のホストマシンには仮想環境として VMware を導入し, 残り一台をパケットキャプチャのためのマシンとした. そして, 4 台のホストマシン上では VMware によってそれぞれ 2 台ずつの仮想マシンを動作させ, この仮想マシン上で Winny を動作させることで P2P トラヒックを取得した. 上記の構成を図 2 に示す.

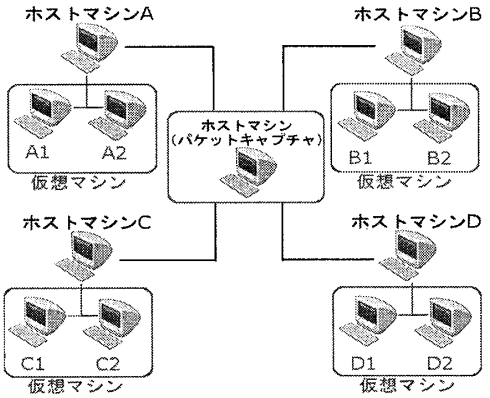

図 2:実験環境の構成

An analytical method for the pure P2P traffic that focus attention on a bidirectional connection for encrypted connection

†Suzuki Masashi, Kyoji Umemura, Hirotake Abe
Information and Computer Science, Toyohashi University of Technology

‡Masayuki Okabe
Information and Media Center, Toyohashi University of Technology

Copyright © 2010 Information Processing Society of Japan.
All Rights Reserved.

3.4 P2P トラヒックの弁別結果

取得した P2P トラヒックから本稿の弁別手法を評価する．今回の実験では α=0.042, β=0.048 とし，P2P トラヒックを用いてフォールスネガティブを，nonP2P トラヒックを用いてフォールスポジティブを検証した．結果を図 3 に示す．

図 3 からフォールスネガティブレート(FNR)は 4%であり，低く抑えられている．また，フォールスポジティブレート(FPR)についても約 0.4%と低く，誤検出の危険性は非常に低いと考えられる．そのため，この弁別手法により P2P トラヒックを正確に弁別できることが期待できる．

トラヒック	コネクション数		FNR[%]
	検出数	合計数	
Winny	5977	6234	4.1226

トラヒック	コネクション数		FPR[%]
	検出数	合計数	
NonP2P	9	2306	0.3903

図 3：フォールスポジティブレート(FPR)と
フォールスネガティブレート(FNR)

4. 通信の暗号化手法

暗号通信が行われる場合に使用される暗号プロトコルとして，IPsec[6]を想定する．IPsec は複数の暗号方式を使用することができるので，通信相手との間で通信の設定を合わせるためのパラメータを共有する必要がある．このとき通信相手と共有するパラメータを SA とよぶ．SA にはさまざまなパラメータがあるが，重要なパラメータとなるのが，セキュリティプロトコルとモードである．セキュリティプロトコルは ESP と AS がある[7]．AH は発信元の認証，完全性認証を提供し，ESP はパケットの暗号化機能を提供する．モードはトランスポートモードとトンネルモードの 2 つがある．トランスポートモードはペイロードだけを，トンネルモードは IP パケット全体をカプセル化する．

本稿ではセキュリティプロトコルに ESP，モードにトランスポートモードを選択して議論を進める．なお，トランスポートモードを選択したことで一般性が失われることはない．

5. 暗号化通信における弁別手法の問題

5.1 IPsec によるパケット長の変更の問題

IPsec によって暗号化を行う場合，パケットに対して暗号化のためにデータが付加される．このためトラヒック弁別のために必要となる特徴量であるパケット長が変化してしまうといった問題がある．

特にパケット長を変化させる大きな要因としてパディングが問題となる．通常のパディングは ESP トレイラに付加され，平文の長さを暗号化アルゴリズムで定められたブロック長の整数倍にするために用いられる．しかし，上記のパディング以外にもデータ部分に付加される TFC(Traffic Flow Confidentiality)パディングが存在する．TFC パディングは 0～255byte の間で設定することができ，データの長さを隠蔽することや，ダミーの ESP パケットを送信してトラヒックの特性を隠蔽するために利用される．また TFC パディングによるダミーパケットは暗号化されてしまっているため，パケット解析によってダミーパケットかどうか判断することはできない．このため，TFC パディングの設定によっては，P2P トラヒックを正確に弁別することが困難となる．

5.2 暗号化処理による遅延の影響

暗号化通信を行う際には暗号化や復号化処理が必要となる．この処理によって 3 ウェイハンドシェイクの SYN パケットを送ってから SYN/ACK パケットが帰ってくるまでの時間に遅延が発生してしまい，P2P トラヒックの弁別に対して影響がでる可能性がある．

6. 暗号化通信における問題点の検証

IPsec によって暗号化されたトラヒックを解析するためにピュア P2P と同様の双方向通信を行うプログラムを作成した．作成したプログラムによって暗号化しているトラヒックと暗号化していないトラヒックに同様の動作をさせ，これらのトラヒックを比較することで IPsec 通信の調査を行った．

6.1 パケット長の増加

すべてのパケット長が 32byte 増加していることを確認した．そのため SYN パケット，SYN/ACK パケット共に 80byte となった．これにより，プロトコルが ESP の際にはパケット長を 80byte に変更することで P2P トラヒックを弁別できると考えられる．

6.2 暗号化処理の遅延

暗号化された P2P トラヒックと暗号化されていない P2P トラヒックから 1000 件の 3 ウェイハンドシェイクを抽出し，それぞれの時間間隔と比較した．結果を図 4 に示す．図 4 から暗号処理による平均の遅延時間は約 0.001 秒であることがわかった．

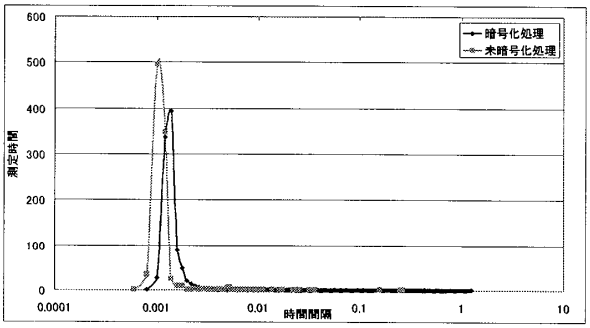

図 4：3 ウェイハンドシェイクの時間間隔比較

7. まとめ

本稿の解析手法によって暗号化された P2P トラヒックを解析する際の問題点を挙げ，その影響を調査した．IPsec によって通信が暗号化された場合にパケットに付加されるデータは 32byte であった．また，暗号化処理による遅延の影響は 0.001 秒ほどであることがわかった．これにより IPsec により暗号化された場合パケット長を 32byte 増加することと，処理の遅延が 0.001 であることを考慮する必要がある．

今後は IPsec によって暗号化された P2P トラヒックに対して弁別手法を適用したいと考えている．また，Winny 以外の P2P ファイル共有アプリケーションに対して本稿の弁別手法が適用できるかどうか検討しようと考えている．

参考文献

[1]三浦明日香，梅村恭司，阿部洋丈，岡部正幸 ：SYN パケットの呼応に着目した P2P トラヒックの表示，情報処理学会全国大会講演論文集，pp.239-240 (2009)

[2]松田崇，中村文隆，若原恭，田中良明 ：相互接続における順逆接続間隔を利用した P2P トラヒック分別手法，信学技報，no.NS2006-237，pp415-420 (2007)

[3]松田崇，中村文隆，若原恭，田中良明，大崎淳，千田浩一，加藤圭，飯塚正 ：PureP2P ファイル共有トラヒックの特性解析，信学技報，no.NS2005-2，pp.5-8 (2005)

[4]" One Point Wall" http://www.onepointwall.jp/

[5]亀井聡 ：P2P 技術がネットワークインフラに及ぼす影響と課題，コンピュータソフトウェア，Vol.22，No.3，pp.8-18，日本ソフトウェア科学会，(2005)

[6]" Security Architecture for the Internet Protocol"，RFC 4301，IETF

[7]" Cryptographic Algorithm Implementation Requirements for Encapsulating Security Payload (ESP) and Authentication Header (AH)"，RFC 4305，IETF

Copyright © 2010 Information Processing Society of Japan.
All Rights Reserved.