平成22年度 情報工学コース卒業研究報告要旨

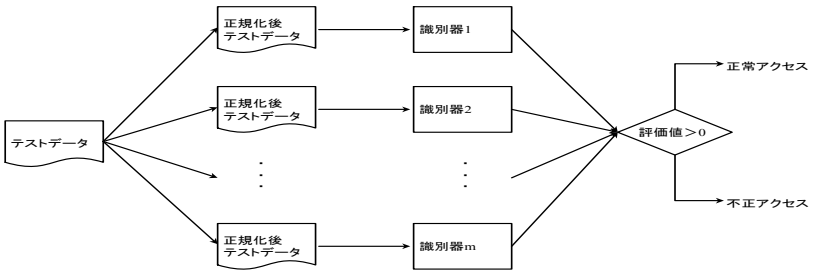
高倉・八槇 研究室	氏　　名	岸 本 和 也
卒業研究題目	期間別識別器の合成による アノマリ型IDSの検知率改善	

近年インターネットを介した攻撃が増加しており，そのような攻撃を防ぐ手法としてIDS(Intrusion Detection System，侵入検知システム) が注目されている．IDS にはあらかじめ定義された不正アクセスパターンと通信データとの照合により攻撃を検知するシグネチャ型と通常時の通信パターンからの逸脱から不正な通信を検知するアノマリ型が存在する．アノマリ型は一般的な検知手法では難しいとされる未知の攻撃なども検知することが可能であるという利点がある．

アノマリ型は機械学習により通信データの識別を行うため，学習データの質により識別率が大きく変化する．インターネットでは，正常な通信や不正な通信が次々と生成される一方で，事前に知られていない未知の攻撃が行われることがあるなど、常に状態が変化し続けている。従って，事前にどのような学習データが適切であるかを予測し，用意するのは困難である．またネットワーク状況の変化に追従しつつ，学習データを常に最適，かつ，最新のものに維持し続けるのはコスト面から見て現実的ではない．

本研究ではアノマリ型IDS の上記の問題を解決するために，ハニーポットや定点観測装置 (Darknet) で観測された通信データを解析し，データのサイズの変化が識別率に及ぼす影響，および，観測日毎のデータの傾向の変化を調査した．その結果，通信の傾向は日々大きく変化するため，単に学習データの期間 (サイズ) を大きくしただけでは過学習に陥り，識別率が低下することを確認した．一方で通信の傾向の変化は極めて大きく，すべての不正アクセスを検知できる代表的な観測日データを見つけることも困難であることが分かった．

本予備調査の結果に基づいて，本研究では，識別率が最善となるサイズの学習データを複数用いて識別機群を作成した．さらにそれらの識別結果を合成することで様々な傾向の通信において高い識別率を示す識別器を作成する手法を開発した．



また刻々と変化する正常通信と不正通信が混在した状況に対応するため，識別機を常に最良の状態に保たねばならない．しかしその状況毎における最適な識別器群を予測することは前述したように困難であり，さらに現在ある識別器群の全てが次の識別において大きく性能を下げるわけではない．そこで新たな識別器を定期的に作成し，合成に使用している識別器群の中から性能の低いものと置き換えることで傾向の変化に追従し，かつ，検知率を高める更新手法を開発した．

本手法の有効性を確認するために，京都大学ハニーポットにより収集された通信データと既存の識別器作成アルゴリズムを用いて性能の評価実験を行った．その結果，検知率が平均 80.93% ，false positive 率が平均 5.90% と安定して高い性能を示した．