論文の内容の要旨

インターネット技術の普及と伴い、近年 Web アプリケーションのデータベースの情報に不正にアクセスする SQL インジェクション攻撃の被害は特に深刻であり、有効な対策が求められる。また、SQL インジェクションに限らず、Web アプリケーション攻撃を防御するための Web アプリケーションファイアウォール(WAF)の開発も急務である。従来の SQL インジェクション攻撃の対策は、エスケープ処理や既存攻撃の正規表現とのパターンマッチングのブラックリスト方式や、機械学習を利用する方法が知られているが、バイパス攻撃の問題、攻撃と構成が似ている正常な入力を誤検知する False Positive の問題、そして今まで知られていない新しい攻撃の対応が問題となっている。

本研究は，SQL インジェクション攻撃の文字列と攻撃の特徴とよく似た正常な文字列の特徴抽出法を開発することで、False Positive と False Negative の両方を小さくし、未知の攻撃に対処できるだけでなく、バイパス攻撃の実現を困難にする WAF を開発することを目的としている。

まず、SQL インジェクション攻撃の文字列中の攻撃特徴記号の確率分布に注目し、その含有率が極めて低いという特徴を記述するために、ベルヌイ分布の拡張を利用した攻撃特徴記号の含有率とその時の攻撃の可能性を表す確率モデルを提案した。また、本モデルのパラメータに対して、対数尤度関数のテーラー展開により理論的に最尤推定量を導出した。この解は Newton-Raphson 法による数値計算によっても確認され、含有率の閾値は、機械学習における SCW と呼ばれる線形分類器からも導かれるため、SQL インジェクション攻撃の検知システムに有効であることが確認された。

次に、SQL インジェクション攻撃に上述の攻撃特徴記号を含む正常な入力の誤判定問題を解決するために、本研究では、潜在曲線モデルを応用した特徴抽出モデルを開発し、攻撃特徴記号同士の関連性を多項式で表現することで、攻撃と正常の両方の特徴を抽出する手法を提案した。特に、攻撃に頻出する記号の上位 5 つの記号を攻撃特徴記号として使用し、攻撃と正常の両方の特徴を表現することができると考えられる 2 次の多項式を用いて特徴抽出をするための潜在曲線モデルを提案した。

提案手法の有効性を検証するために，オープンソースウェア(OSS)の WAF として有名である ModSecurity と、汎化能力が高いことで知られているサポートベクターマシン（SVM）を用いて攻撃検知との比較実験を行なった結果、従来方式より優れた性能が確認できた。なお，本研究で提案したモデルとパラメータ推定法に基づく攻撃検知システムを Apache のモジュールとして開発した。

論文審査の結果の要旨

本研究は、ネットワークセキュリティの攻撃検知において、確率モデルを用いる方法論の樹立と、効率的なWAFシステムの開発を目指したものである。

従来の攻撃検知法は、既知攻撃のシグネチャなどのパターンを記憶して、これらのパターン照合によるブラックリスト方式が主であるが、これらの手法の最大の課題は、過剰検知による正常入力の誤判定と、未知なる攻撃に対する対応能力の欠如である。機械学習による方法も、限られたデータによる学習の汎化能力を頼りに新しい攻撃パターンへの対応は難しい。

本論文は，SQLインジェクション攻撃の固有の特徴を捉える確率モデルを構築することで、未知なる攻撃にも対応可能な攻撃検知方式を提案し、効率的なWAFシステムを開発している。

本論文は、以下のように構成されている。第1章「序論」は研究背景、当分野の課題と研究目的を述べている。第2章「SQLインジェクション攻撃」はSQLインジェクション攻撃を紹介し、具体例を示している。第3章「既存研究のまとめ」は今までSQLインジェクション攻撃に対する主な対策と検知手法を紹介し、それらの利点と課題を分析している。第4章「SQLインジェクション攻撃と記号分布」は、従来手法の課題を解決するために、SQLインジェクション攻撃の文字列における攻撃特徴記号の確率分布を表すモデルを提案し、その最尤推定に基づく攻撃検知法を示している。第5章「潜在曲線モデルによる攻撃検知アルゴリズム」は、さらに記述力の高い潜在曲線モデルを導入し、新しい攻撃検知法を示し、WAFシステムを構築している。第6章「従来研究との比較と考察」は、提案手法と既存手法との比較検討を行っている。第7章「考察」は提案手法のネットワークセキュリティにおける意義と期待される役割を議論している。第8章「結論」は本研究のまとめと今後の課題を述べている。

従来、攻撃パターンを蓄積して照合する方法と機械学習を用いる方法は、攻撃と正常入力を区別するのが難しく、新しい攻撃に弱いという問題に対して、攻撃文と正常文両方の特徴を捉えた複数記号の同時確率分布を表すモデルに基づく確率モデルを構築する手法には、潜在的可能性が秘められていると考えられ、正常入力の誤判定の解消と未知攻撃への対応能力が期待される。

一方、本研究の確率モデルが攻撃を適切に記述しているかどうかが最重要課題であり、モデルの妥当性と選択法については、さらなる検討が不可欠である。

本論文は、情報セキュリティの基礎研究に貢献するのみならず、SQLインジェクション攻撃検知の実用価値の高い方式と効率的なWAFシステムを開発している。科学技術の理論と応用の発展に寄与することが認められ、博士（工学）論文として十分な価値があると考える。