STUDIA INFORMATICA 2003
Volume 24 Number 2A (53)

Ewa LACH
Politechnika Śląska, Instytut Informatyki

SZTUCZNA INTELIGENCJA W SYSTEMACH WYKRYWANIA INTRUZÓW

Streszczenie. W pracy przedstawiono wybrane techniki sztucznej inteligencji wykorzystywane w systemach wykrywania intruzów. Artykuł zawiera opis sieci neuronowych, odkrywania wiedzy, logiki rozmytej, algorytmów genetycznych, programowania genetycznego oraz sztucznych systemów immunologicznych.

Słowa kluczowe: sztuczna inteligencja, systemy wykrywania intruzów.

ARTIFICIAL INTELLIGENCE IN INTRUSION DETECTION SYSTEMS

Summary. The paper survey Artificial Intelligence techniques used in Intrusion Detection systems. Several approaches to intrusion detection were described including: neural networks, data mining, fuzzy logic, genetic algorithms, genetic programming and artificial immune systems.

Keywords: artificial intelligence, intrusion detection.

1. Wprowadzenie

Rosnąca z roku na rok ilość włamań, powszechność sieci lokalnych i Internetu powoduje wzrost zapotrzebowania na systemy służące do monitorowania nielegalnych zachowań. Coraz więcej organizacji wdraża systemy wykrywania intruzów – IDS (ang. Intrusion Detection Systems) identyfikujące akcje naruszające integralność, poufność lub dostępność zasobów.

Najważniejszą cechą systemów wykrywania intruzów jest umiejętność rozróżnienia normalnych, akceptowalnych zachowań użytkowników od zachowań nienormalnych, mogących być próbą ataku. Wyróżniamy dwa podejścia do tego problemu: analizowanie sygnatur (ang. misuse detection) oraz wykrywanie anomalii (ang. anomaly detection).

Systemy IDS stosujące analizę wzorców korzystają z bazy wiedzy o technikach włamań, porównując znane wzorce ataków z monitorowanymi działaniami zachodzącymi w sieci lub systemie. Poważnym minusem tego podejścia jest jego niezdolność do wykrywania zdarzeń, których sygnatury nie zostały umieszczone w bazie.

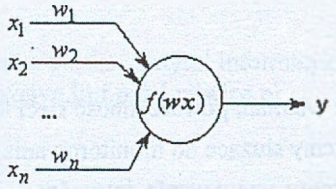
Przy wykrywaniu anomalii system IDS bazuje na profilu oczekiwanego zachowania systemu lub użytkowników. Jakiekolwiek znaczące odstępstwo od tego profilu jest traktowane jako potencjalne włamanie. Podstawową zaletą tej metody jest jej zdolność do wykrywania nowych ataków, dla których wzorce nie zostały jeszcze zdefiniowane. Istotną wadą tego podejścia jest otrzymywanie dużej liczby alarmów odnoszących się do legalnych działań użytkowników, które system klasyfikuje jako potencjalnie niebezpieczne.

2. Wybrane metody sztucznej inteligencji

Systemy wykrywania intruzów z ich koniecznością przejrzenia i zaklasyfikowania ogromnej ilości informacji wykazują istotne zapotrzebowanie na rozwiązania powstałe w ramach ogólnie pojmowanej sztucznej inteligencji. Obecnie trwają badania nad zastosowaniem w systemach IDS sieci neuronowych, algorytmów genetycznych, programowania genetycznego, logiki rozmytej, odkrywania wiedzy, systemów immunologicznych i innych technik korzystających ze zdobyczy sztucznej inteligencji.

2.1. Sztuczne sieci neuronowe

Sztuczna sieć neuronowa symuluje pracę mózgu. Składa się ona z neuronów i połączeń między nimi, do których są przypisane określone wagi wskazujące na siłę połączenia. Poniższy rysunek przedstawia schemat najczęściej stosowanego modelu sztucznego neuronu.



Rys. 1. Schemat modelu sztucznego neuronu
Fig. 1. Diagram of artificial neuron

Wartość sygnału wyjściowego y neuronu jest określona poprzez następującą relację:

$$y = f(wx) \tag{1}$$

gdzie w jest wektorem wag, x jest wektorem wartości sygnałów wejściowych, natomiast funkcja f jest to funkcją aktywacji neuronu.

Poszczególne jednostki neuronowe łączą się tworząc sieć neuronową. Obecne systemy IDS wykorzystują dwa rodzaje topologii sieci: jednokierunkowe i rekurencyjne. Sieci jednokierunkowe, stosowane w większości przypadków, tworzone są w wyniku grupowania pojedynczych neuronów w warstwy, a następnie łączenia warstw w ten sposób, że wyjście każdego neuronu z danej warstwy jest połączone z wejściem każdego neuronu z warstwy następnej. Nie mogą istnieć połączenia wewnątrz warstwy oraz połączenia wsteczne. W sieciach rekurencyjnych wyjście przynajmniej jednego neuronu jest połączone pośrednio lub bezpośrednio z jego wejściem.

Manipulując wagami możemy nauczyć sieci neuronowe rozwiązywania różnych skomplikowanych problemów.

Wyróżniamy kilka kryteriów klasyfikacji sieci neuronowych. Z punktu widzenia systemów IDS najistotniejszy jest podział ze względu na tryb uczenia sieci:

- uczenie nadzorowane (z nauczycielem) – wykorzystuje się wcześniej stworzone dane uczące składające się z informacji wejściowych i odpowiadających im poprawnych odpowiedzi (informacji wyjściowych). Nauka sieci neuronowej polega na dopasowaniu wag, tak aby dane wejściowe były klasyfikowane poprawnie,
- uczenie nienadzorowane – sieci neuronowe starają się znaleźć wzorce w otrzymanych na wejściu danych, które klasyfikują według najistotniejszych cech.

W systemach IDS sieci neuronowe biorą udział w tworzeniu modeli wzorców zachowań, a następnie w klasyfikacji nowych sygnatur jako akceptowalnych lub nielegalnych.

Najbardziej popularną siecią stosowaną w systemach IDS jest wielowarstwowa sieć jednokierunkowa z algorytmem wstecznej propagacji błędów (ang. back-propagation). Sygnały dostarczone do warstwy wejściowej są przetwarzane w kolejnych warstwach ukrytych, aż dochodzą do warstwy wyjściowej. Funkcja aktywacji neuronu jest dowolna. W trakcie przejścia wstecznego rzeczywisty sygnał wyjściowy jest porównywany z oczekiwanym i obliczany jest szacunkowy błąd dla wyjściowej jednostki. Następnie wagi odnoszące się do danej jednostki są dostosowywane, aby zminimalizować ten błąd. Posuwając się wstecz uzgadniamy kolejne wagi w warstwach ukrytych, aż dochodzimy do warstwy wejściowej.

Systemy IDS oparte na powyższym modelu zostały przedstawione w [1, 2, 3, 4, 5] i różnią się między sobą głównie dziedziną danych, które monitorują. Wczesne prace nad systemami wykrywania intruzów używających sieci neuronowych skupiały się przede wszystkim na wykrywaniu anomalii, ucząc sieci neuronowe w oparciu o normalne zachowanie sieci i użytkowników. Tylko część autorów stosowała sieci neuronowe do wykrywania sygnatur ataków. W tym wypadku dążono do generalizacji zachowań

nielegalnych, w przeciwieństwie do tworzenia szczegółowych sygnatur dla poszczególnych ataków.

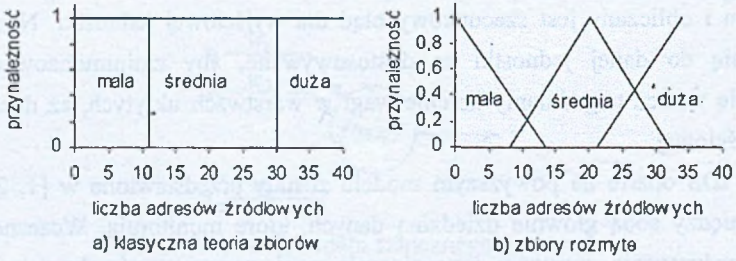
W ostatnich badaniach nad systemami IDS skupiono się nad modelami nienadzorowanego uczenia, bazującymi na samoorganizujących się sieciach (ang. self-organizing maps) - SOM. Samoorganizujące się sieci oparte są na uczeniu konkurencyjnym, które klasyfikuje wzorce wejściowe według kryterium podobieństwa. Techniki te zostały zastosowane w analizie zachowań użytkowników, aplikacji i procesów ([6, 7, 8]).

2.2. Odkrywanie wiedzy. Logika rozmyta

Odkrywanie wiedzy (ang.: Data Mining), dotyczące automatycznego poszukiwania ukrytych, nieznanych wcześniej reguł w dużych zbiorach danych, znalazło istotne zastosowanie w IDS. W przypadku systemów analizy sygnatur znanych ataków, odkrywanie wiedzy umożliwiło zautomatyzowanie niezwykle żmudnego i podatnego na błędy procesu tworzenia reguł. Przykładowym systemem wykorzystującym odkrywanie wiedzy do tworzenia wzorców ataków jest MADAM ID (Mining Audit Data for Automated Models for Intusion Detection) utworzony na Uniwersytecie Columbia ([13]). Z odkrywania wiedzy skorzystano także przy tworzeniu wzorców zachowań oczekiwanych. Ta technika została zastosowana np. w systemie ADAM (Audit Data Analysis and Mining) opisanym w [14] oraz IDDM (Intusion Detection using Data Mining) wyjaśnionym w [15].

Kolejnym kierunkiem badań jest zaproponowana przez Susan Bridges ([16]) integracja odkrywania wiedzy z logiką rozmytą, zapewniająca większą ogólność i elastyczność reguł.

Logika rozmyta uogólnia klasyczną Cantorowską teorię zbiorów, która zakłada, że dowolny element należy lub nie należy do danego zbioru. W teorii zbiorów rozmytych element może tylko częściowo należeć do pewnego zbioru.



Rys. 2. Przedstawienie zmiennych jakościowych w logice klasycznej i rozmytej
Fig. 2. Non-fuzzy and fuzzy representations of sets for quantitative variables

Załóżmy, że mamy prostą regułę określającą podejrzaną sytuację:

„Jeżeli liczba różnych adresów źródłowych w przeciągu ostatnich 2 sekund jest wysoka, oznacza to, że mamy do czynienia z nielegalnym zachowaniem."

Powyższa reguła otrzymana dzięki odkrywaniu wiedzy mogła by mieć następującą postać:

IF s=2 **AND** s_address > 30 **THEN** wynik = atak

Używając logiki tradycyjnej należy podać punkt, powyżej którego liczba jest duża. Powoduje to, że niewielka zmiana wartości zmiennej (w naszym przypadku przejście z 30 do 31) może spowodować istotną zmianę klasyfikacji sytuacji (rys.2a.).

W przypadku reguł rozmytych elementy należą do zbiorów z danym stopniem przynależności (z zakresu [0,1]) określonym przez funkcję przynależności.

Jak zostało pokazane na rysunku 2b dla naszego problemu, 30 adresów w stopniu 0,7 należy do zbioru duża, a w stopniu 0,2 do zbioru średnia.

Wykorzystując logikę rozmytą, powyższej regule moglibyśmy nadać następującą postać:

IF s=2 **AND** s_address = duża **THEN** wynik=atak

Wielkość udziału zmiennej *s_address* w zbiorze *duża* miałaby następnie wpływ na określenie wielkości zaufania dla danej reguły. Im większe zaufanie, tym bardziej jest prawdopodobne, że dana reguła dołączy do modelu sygnatur.

2.3. Programowanie ewolucyjne

Algorytmy ewolucyjne jest to klasa algorytmów heurystycznych naśladujących procesy naturalne: dziedziczenie genetyczne i darwinowską walkę o przeżycie. W programowaniu ewolucyjnym występuje populacja struktur (osobników), którym przypisuje się miarę ich dopasowania do otoczenia. Wybierając do następnego pokolenia osobniki najlepiej przystosowane (operacja selekcji) zapewnia się, że każda kolejna generacja jest średnio lepiej przystosowana niż poprzednia. Dokonując określonych przez operatory genetyczne działań na osobnikach zapewnia się eksplorację potencjalnych rozwiązań.

W systemach IDS wykorzystuje się przede wszystkim dwie techniki programowania ewolucyjnego: algorytmy genetyczne oraz programowanie genetyczne. Podstawowa różnica między nimi dotyczy osobników populacji. Osobnikami w algorytmach genetycznych są ciągi o ustalonej długości, będące rozwiązaniami zadania, podczas gdy w programowaniu genetycznym osobniki stanowią programy, które powinny dostarczyć rozwiązanie problemu.

Algorytmy genetyczne znalazły zastosowanie w systemach wykrywania intruzów przy generowaniu i adaptowaniu do zmieniającego się otoczenia reguł klasyfikatora zdarzeń (przykładowy projekt opisano w [17]). Wykorzystano je także do dostrajania wartości oraz funkcji przyjmowanych z góry przez inne techniki (np. wybór funkcji przynależności w logice rozmytej, ustalanie początkowych wag dla sieci neuronowych) oraz do wyboru cech, które mają największy wpływ przy klasyfikowaniu zdarzeń.

Programowanie genetyczne zastosowano do generowania i ewoluowania agentów, stanowiących niezależne systemy IDS (np.[18]).

2.4. Sztuczne systemy immunologiczne

W ostatnich latach pojawiła się duża ilość badań nad systemami wykrywania intruzów zainspirowanymi systemami immunologicznymi organizmów, umożliwiającymi przetwarzanie równoległe oparte na niezależnych agentach. Badane systemy wykorzystują fakt, że układy obronne organizmów mają dużo cech systemów wykrywania anomalii (wiedzę o tym, co jest normalne i umiejętność identyfikacji tego, co jest różne od normalnego) oraz istotne cechy systemów analizy sygnatur (umiejętność identyfikacji wzorców, pamięć o przeszłych atakach, lepsze wykrywane znanych włamań).

Proces wykrywania anomalii korzysta z algorytmu negatywnej selekcji, który składa się z trzech podstawowych kroków:

- Utworzenie zbioru S reprezentującego własne epitopy (zachowania normalne)
- Utworzenie zbioru receptorów R, które nie rozpoznają żadnego elementu $s \in S$. Receptory są tworzone losowo, następnie za pomocą reguł częściowego dopasowania, receptory reagujące na własne epitopy są eliminowane.
- Monitorowanie nowych zachowań w systemie polegające na nieustannym dopasowywaniu receptorów do wzorców bieżących zachowań. O pojawieniu się intruzów w systemie informuje fakt dopasowania się receptorów.

Aby zmniejszyć ilość fałszywych alarmów, część systemów IDS korzysta z kolejnej właściwości układów odpornościowych: aby został uruchomiony proces niszczenia obcych komórek, odpowiednia liczba receptorów musi zawiązać obce epitopy; ponadto wiązanie receptor - epitop nie trwa wiecznie, co oznacza, że odpowiednia liczba receptorów musi być związana w stosunkowo krótkim czasie. Przykładowo w systemie ARTIS, przedstawionym w [9], przyjęto regułę, że zanim detektor wyśle sygnał o wykryciu anomalii, musi rozpoznać τ epitopów w ustalonym przedziale czasowym.

W układzie odpornościowym uaktywnione komórki, które rozpoznały atak, podlegają klonowaniu. Dodatkowo klony podlegają intensywnej mutacji, tzw.hipermutacji somatycznej, której celem jest wyprodukowanie receptorów możliwie najlepiej dopasowanych do epitopów komórek atakujących. Klony słabo dopasowane są usuwane, natomiast klony dobrze dopasowane przeżywają i po pewnym czasie zmieniają się w komórki pamięciowe.

Systemy wykrywania intruzów naśladują powyższe zachowanie tworząc klony receptorów, które rozpoznały intruzów i przesyłają je do innych węzłów monitorowanej sieci.

Kolejną cechą systemu immunologicznego, znajdującą zastosowanie w systemach IDS, jest fakt, że limfocyty odpowiedzialne za rozpoznawanie obcych epitopów żyją krótko (kilka dni), a ich miejsce zajmuje nowe pokolenie. W systemach wykrywania intruzów wykorzystanie tej właściwości zapewnia dynamiczną adaptację systemu.

Opierając się na systemach immunologicznych IDS może nie tylko wykrywać anomalie, ale także na podstawie wykrytych ataków może tworzyć sygnatury zapamiętywane w bazie wzorców. Taka baza jest następnie wykorzystywana do wykrywania znanych ataków, co przyspiesza rozpoznanie i reakcje na włamanie. Przykładowym systemem IDS, który implementuje to hybrydowe podejście, jest AdenoIdS zaprezentowany w [10].

3. Podsumowanie

Obecnie prowadzone są intensywne badania nad przedstawionymi w poprzednim punkcie technikami. Metody sztucznej inteligencji zmniejszają ilość pracy wymaganej od twórców systemów IDS oraz potrafią poprawić działanie tych systemów. Jest to jednak dziedzina, w której trzeba jeszcze wiele zrobić. Większość z dotychczasowych rozwiązań charakteryzuje się dużą liczbą fałszywych alarmów, długim czasem wdrożenia oraz znacznie mniejszą skutecznością w porównaniu z występującymi obecnie na rynku systemami analizy sygnatur.

LITERATURA

1. Lin M., Mikkulainen R., Ryan J.: Intrusion Detection with Neural Networks. MIT, 1998.
2. Zhang Z., Li J., Manikopoulos C. N., Jorgenson J., Ucles J.: A Hierarchical Network Intrusion Detection System Using Statistical Preprocessing and Neural Network Classification. IEEE Information Assurance Workshop, 2001.
3. Ghosh A., Schwartzbard A.: A Study in Using Neural Networks for Anomaly and Misuse Detection. http://www.usenix.org/events/sec99/ghost.html, 1999.
4. Cannady J.: Artificial Neural Networks for Misuse Detection. NISSC, 1998.
5. Mukkamala S., Janoski G., Sung A.: Intrusion Detection Using Neural Networks and Support Vector Machines. IEEE IJCNN, 2002.
6. Lichodzijewski P., Zincir-Heywood A., Heywood M.: Dynamic Intrusion Detection Using Self Organizing Maps. Canadian Information Technology Security Symposium, 2002.
7. Liu Z., Florez G., Bridges S. M.: A Comparison Of Input Representations In Neural Networks: A Case Study In Intrusion Detection. IJCNN, 2002.
8. Dipankar D., Hal B.: Mobile Security Agents for Network Traffic Analysis. IEEE Computer Society Press, 2001.
9. Hofmeyr S. A., Forrest S.: Architecture for an artificial immune system. EC, 2000.
10. Paula F. S., Reis M. A., Fernandes D. A. M., Geus P. L.: ADenoIdS: A hybrid IDS based on the immune system. ICONIP, 2002.

11. Dasgupta D., Majumdar N. S.: Anomaly detection in multidimensional data using negative selection algorithm. IEEE, 2002.
12. Wierzchoń S. T.: Sztuczne systemy immunologiczne. EXIT, Warszwa 2001.
13. Lee W., Tolfo J., Mok K. W.: A data mining framework for building intrusion detection models. IEEE Symposium on Security and Privacy, 1999.
14. Couto J., Jajodia S., Wu N.: ADAM: Detecting Intrusions by Data Mining. IEEE, 2001.
15. Abraham T.: IDDM: Intrusion Detection using Data Mining Techniques. Technical Report DSTO-GD-0286, 2001.
16. Florez G., Bridges S. M., Vaughn R. B.: An Improved Algorithm for Fuzzy Data Mining for Intrusion Detection. NAFIPS, 2002.
17. Dasgupta D., Gonzalez F.A.: An Intelligent Decision Support System for Intrusion Detection and Response. MMM-ACNS, 2001.
18. Crosbie M., Spafford G.: Applying Genetic Programming to Intrusion Detection. Prude University 1995.

Recenzent: Prof. dr hab. inż. Konrad Wojciechowski

Wpłynęło do Redakcji 1 kwietnia 2003 r.

Abstract

With the tremendous growth of accessibility to the Internet, the prevention and detection of intrusion and misuses has become an issue of serious global concern. Intrusion Detection systems offer techniques for modeling and recognizing normal and abusive system behaviour. There are researches on using Artificial Intelligence methods in IDS systems, which can reduce the human effort required to build these systems and can improve their performance.

This paper survey Artificial Intelligence techniques used in Intrusion Detection systems. Several approaches to intrusion detection were described including: neural networks, data mining, fuzzy logic, genetic algorithms, genetic programming and artificial immune systems.

Adres

Ewa LACH: Politechnika Śląska, Instytut Informatyki, ul. Akademicka 16, 44-101 Gliwice, Polska, elach@iinf.polsl.gliwice.pl .