# WIRELESS INTRUSION DETECTION SYSTEM BASED ON DATA MINING

**Radovan Dvorský**
Master Degree Programme (3), FIT BUT
E-mail: xdvors08@stud.fit.vutbr.cz

Supervised by: Matej Kačic
E-mail: ikacic@fit.vutbr.cz

**Abstract**: Widespread use of wireless networks has made security a serious issue. This paper proposes misuse based intrusion detection system for wireless networks, which applies artificial neural network to captured frames for purpose of anomalous patterns recognition. To address the problem of high positive alarm rate, this paper presents a method of applying two artificial neural networks.

**Keywords**: Intrusion detection system, misuse detection, wireless security, artificial neural network

## 1 ÚVOD

V posledných rokoch sme svedkami veľkého nárastu popularity bezdrôtových sietí. Medzi najznámejšie a najpoužívanejšie patria technológie štandardu IEEE 802.11, známe ako Wi-Fi. Dnes tento štandard podporuje prakticky každé mobilné zariadenie a umožňuje tak jednoduché pripojenie nielen do domácej, ale aj do podnikovej siete. Tento rýchly nárast na popularite mal však negatívne dôsledky v oblasti bezpečnosti, kde aj napriek neustálemu vývoju štandardu a jeho mechanizmov zabezpečenia, stále existujú hrozby umožňujúce útočníkom preniknutie do siete. Odpoveďou na tieto hrozby sú práve detekčné systémy, ktorých úlohou je útoky detegovať a zabrániť ohrozeniu siete. Existuje mnoho metód a prístupov k detekcii, všetky však možno rozdeliť do dvoch základných skupín. Detekčné systémy založené na detekcii anomálií a detekčné systémy založené na hľadaní signatúr. Práve návrhom a popisom detekčného systému založeného hľadaní signatúr sa zaoberá tento článok.

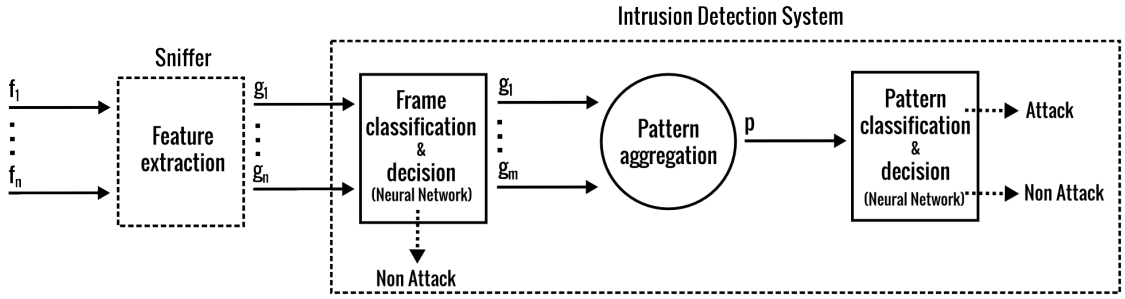
## 2 SÚVISIACE PRÁCE

V práci [1] popisujú autori detekčný systém pre Wi-Fi siete využívajúci viacvrstvovú (MLP) neurónovú sieť so spätným šírením chyby. Práca ukázala, že riešenie pomocou neurónovej siete poskytuje v porovnaní s inými detekčnými metódami [2, 3] lepšie výsledky detekcie s nízkou mierou falošných poplachov. Ďalší prístup k detekcii predstavuje práca [4], v ktorej je popísaný dvojvrstvový detekčný systém pre TCP/IP siete využívajúci na klasifikáciu skupiny udalostí Kohonenovu samorganizujúcu mapu a MLP sieť. Výhodou tohto riešenia je možnosť identifikovať komplexné vzory útokov vyskytujúce sa v dlhšom časovom horizonte. Pri tvorbe a návrhu detekčného systému boli využité práve výsledky týchto prác, kde hlavným cieľom bolo vytvoriť detekčný systém, ktorý by dosahoval vysokej úspešnosti s minimálnym množstvom falošných poplachov pri čo najširšom pokrytí útokov.

## 3 ARCHITEKTÚRA APLIKÁCIE

Celý detekčný systém pozostáva z dvoch komponent. Prvou je sniffer, resp. senzor, ktorého úlohou je extrakcia atribútov z hlavičiek rámcov a ich následné uloženie do databázy. Pre potreby vytvárania testovacích dát bolo ďalej nutné zabezpečiť identifikáciu a označenie dátovej komunikácie v prípade útoku. To bolo dosiahnuté začlenením vykonávania externých skriptov (útokov) priamo do sniffera,

čo umožnilo presne vymedziť začiatok a koniec útoku a súčasne aj automatizovať celý proces tvorby testovacích dát. Druhou komponentou je samotný detekčný systém zobrazený na obrázku č. 1.




**Obr. 1:** Schéma architektúry detekčného systému.

V prvom kroku sniffer z hlavičiek rámcov extrahuje 12 atribútov, ktoré spolu tvoria vstupný vektor prvej neurónovej siete. Jej úlohou je klasifikovať rámce na normálne a potencionálne útočné rámce, ktoré sú v ďalšom kroku agregované do vzorov dĺžky $n$. V tomto kroku sa počítajú charakteristiky takto vytvorených vzorov (napr. pomer autentifikačných rámcov k celkovému počtu rámcov), ktorých je spolu 18 a tvoria vstupný vektor druhej neurónovej siete. Druhá neurónová sieť má potom za úlohou klasifikovať tieto charakteristiky a následne ná základe určenej „threshold“ hodnoty rozhodnúť či sa jedná o útok alebo ide o normálnu dátovú komunikáciu.

## 4 METODIKA TESTOVANIA

Pri testovaní boli v detekčnom procese použité dve plne prepojené MLP siete s aktivačnou funkciou v tvare $Sig(x) = \frac{1}{1-e^{-x}}$. V prvom prípade mala sieť vo vstupnej vrstve spolu 12 neurónov (pre každý atribút jeden), nasledovali 3 skryté vrstvy a vo výstupnej vrstve 6 neurónov (pre každý útok jeden), ktoré v závislosti od útoku nadobúdali dvoch hodnôt (0 a 1). Druhá sieť mala podobnú topológiu, jej vstupná vrstva obsahovala 18 neurónov, 5 skrytých vrstiev a 6 neurónov vo výstupnej vrstve. Počet skrytých vrstiev a uzlov v nich bol v obidvoch prípadoch stanovený metódou pokus-omyl, s cieľom dosiahnuť minimálnej chyby pri trénovaní sietí. Ako dĺžka vzoru $n$ bola zvolená hodnota 180 rámcov.

Následne boli pomocou sniffera vytvorené trénovacie a testovacie dáta. Trénovacie dáta pozostávali z bežnej dátovej komunikácie a útokov zbieraných v rámci jedného sedenia. Celá kolekcia obsahovala približne 500 tis. rámcov, z čoho zhruba 200 tis. tvorili útočné rámce. Testovacie dáta boli vytvorené podobný spôsobom, s tým rozdielom, že táto kolekcia bola vytváraná v rámci niekoľkých sedení v priebehu dvoch dní. Kolekcia navyše obsahovala aj dátovú komunikáciu okolitých prístupových bodov a celkovo pozostávala z približne 1 mil. rámcov, z čoho bolo zhruba 300 tis. útočných. Na takto vytvorených testovacích dátach bola následne vykonaná validácia natrénovaného modelu.

## 5 DOSIAHNUTÉ VÝSLEDKY

Výsledky testovania ukazujú tabuľky 1 a 2, kde tabuľka č. 1 zobrazuje úspešnosť klasifikácie rámcov a tabuľka č. 2 úspešnosť detekcie po klasifikácii vzorov. Z prvej tabuľky je vidieť, že klasifikátory rámcov dosiahli dobrej celkovej úspešnosti blížiacej sa k 100 %. Rovnako bolo u väčšiny útokov dosiahnutej dobrej hodnoty „hit“ uvádzajúcej percento v ktorej klasifikátor správne predpovedal útok pri jeho výskyte. V prípade niektorých útokov boli však podľa predpokladu dosiahnuté vyššie „false positive“ hodnoty. Z druhej tabuľky je však vidieť výrazné zníženie „false positive“ klasifikácií, kde s výnimkou *Auth. Flood* útoku došlo k zníženiu až k hodnotám blížiacim sa 0 %. Rovnako boli

dosiahnuté dobré „hit“ hodnoty, kde došlo len k malému poklesu v prípade *KoreK ChopChop* a *Fragmentation* útoku. Z celkového hľadiska sa však miera úspešnosti pri všetkých útokoch zachovala.

| Útok | False positive | False negative | Hit | Úspešnosť |
|---|---|---|---|---|
| Authentication Flood | 1.44 % | 0 % | 100 % | 98.56 % |
| Deauthentication Flood | 0.03 % | 0.03 % | 98.87 % | 99.94 % |
| CTS Flood | 0.37 % | 0 % | 100 % | 99.63 % |
| Korek ChopChop | 3.99 % | 0 % | 100 % | 96.01 % |
| Fragmentation útok | 3.07 % | 0.1 % | 97.62 % | 96.92 % |
| ARP replay | 1.43 % | 0 % | 99.99 % | 98.57 % |

**Tabuľka 1:** Úspešnosť klasifikácie rámcov.

| Útok | False positive | False negative | Hit | Úspešnosť |
|---|---|---|---|---|
| Authentication Flood | 0.96 % | 0.05 % | 99.80% | 98.99 % |
| Deauthentication Flood | 0 % | 0.06 % | 98.57 % | 99.94 % |
| CTS Flood | 0.04 % | 0 % | 100 % | 99.96 % |
| Korek ChopChop | 0.14 % | 0.18 % | 91.49 % | 99.66 % |
| Fragmentation útok | 0 % | 0.05 % | 91.67 % | 99.95 % |
| ARP replay | 0.24 % | 0.05 % | 99.66 % | 99.71 % |

**Tabuľka 2:** Úspešnosť klasifikácie vzorov.

## 6 ZÁVER

V práci bol predstavený detekčný systém a jeho architektúra využívajúca dve neurónové siete. Podarilo sa vytvoriť jeho prototyp a pre potreby testovania bol vytvorený jednoduchý systém automatizovaného vytvárania testovacích dát. Pomocou týchto dát bola v reálnej prevádzke overená úspešnosť navrhnutého detekčného systému, pričom bola dosiahnutá vysoká miera úspešnosti detekcie pri nízkom počte falošných poplachov. V ďalšej práci bude treba systém rozšíriť o ďalšie útoky a taktiež bude treba skúmať atribúty Wi-Fi komunikácie a navrhnúť metriky zvyšujúce úspešnosť detekcie.

## LITERATÚRA


[1] Ezeife, C. I.; Rahman, M. Z.: NeuDetect: A Neural Network Data Mining Wireless Network Intrusion Detection System. In *Proceedings of the Fourteenth International Database Engineering & Applications Symposium*, IDEAS '10, New York, NY, USA: ACM, 2010, ISBN 978-1-60558-900-8, s. 38–41.

[2] Rahman, A.; Ezeife, C. I.; Aggarwal, A. K.: WiFi Miner: An Online Apriori-Infrequent Based Wireless Intrusion System. In *KDD Workshop on Knowledge Discovery from Sensor Data*, *Lecture Notes in Computer Science*, ročník 5840, Springer, 2008, ISBN 978-3-642-12518-8.

[3] Ezeife, C. I.; Ejelike, M.; Aggarwal, A. K.: WIDS: a sensor-based online mining wireless intrusion detection system. In *IDEAS*, *ACM International Conference Proceeding Series*, ročník 299, editácia B. C. Desai, ACM, 2008, ISBN 978-1-60558-188-0, s. 255–261.

[4] Cannady, J.: Artificial neural networks for misuse detection. In *National Information Systems Security Conference*, 1998.