Jerzy Konorski
Politechnika Gdańska, Wydział ETI
ul. Narutowicza 11/12, 80-233 Gdańsk
jekon@eti.pg.gda.pl

KSTI 2017 Wojskowa Akademia Techniczna
Warszawa, 13-15 września 2017

MODELOWANIE I OGRANICZANIE SKUTKÓW ATAKÓW UZURPACJI UPRAWNIEŃ W SYSTEMACH TELEINFORMATYCZNYCH WSPIERAJĄCYCH RÓŻNICOWANIE POZIOMU QOS

MODELING AND MITIGATING FAKE VIP ATTACKS IN COMPUTER COMMUNICATION SYSTEMS SUPPORTING QOS DIFFERENTIATION

Streszczenie: **W wieloagentowych systemach teleinformatycznych pracujących w paradygmacie Klient-Serwer i wspierających różnicowanie poziomu QoS poważnym zagrożeniem są ataki uzurpacji uprawnień metodą *Fałszywego VIPa*. Ich celem jest zapewnienie Klientowi nienależnie wysokiego poziomu QoS, co powoduje nadużycie zasobów Serwera i szkody dla innych Klientów. W referacie podjęto próbę sformalizowania modelu takich ataków oraz zaproponowano *obustronnie ślepy* podsystem reputacyjny ograniczający ich skutki przy minimalistycznym modelu wiedzy zaangażowanych agentów.**

Abstract: **A serious threat to multi-agent computer communication systems offering QoS differentiation in a client-server paradigm are *fake VIP attacks* that consist in falsely declaring a high class of request to acquire undue service quality. This may deplete the server's resources and damage other clients' QoS. The paper proposes a formal approach to such attacks and a double-blind reputation scheme to mitigate their effects under a minimalist information model of the involved agents.**

Słowa kluczowe**: jakość usług, Fałszywy VIP, detekcja sygnatur, reputacja, zaufanie.**

Keywords**: QoS, Fake VIP attack, signature detection, reputation, trust.**

1. WSTĘP

W wieloagentowych systemach teleinformatycznych interakcje między agentami zachodzą w paradygmacie Klient-Serwer: Klient żąda określonego poziomu jakości usług (ang. *quality of service*, QoS), zaś Serwer zapewnia wymagany QoS. Dla odzwierciedlenia działania dzisiejszych złożonych systemów, takich jak systemy chmurowe, wielousługowe sieci pakietowe, portale *e-commerce*, farmy serwerów, rozproszone bazy danych czy aplikacje *online*, paradygmat ten można wzbogacić o moduły Nadawcy i Odbiorcy. Klient generuje abstrakcyjne *obiekty* (np. pakiety, zapytania lub transakcje) różnych *klas natywnych*. Klasa natywna jest właściwością obiektu uprawniającą do określonego poziomu jakości usług. Nadawca, działając na rzecz Klienta, dołącza do każdego obiektu zależną od klasy natywnej informację nazywaną *klasą żądaną* (np. nagłówek, flagę lub odpowiednie metadane) i przekazuje obiekt do Odbiorcy jako żądanie usług. Odbiorca decyduje o *klasie przydzielonej* (tj. QoS) dla obiektu, działając na rzecz Serwera i ochrony jego zasobów, toteż klasa przydzielona może różnić się od klasy żądanej; następnie obiekt przekazany zostaje do Serwera, który udostępnia QoS zgodnie z klasą przydzieloną. Odbiorca jest więc modelem np. tranzytowego węzła sieciowego z działającym mechanizmem routingu lub równoważenia obciążenia, procesora czołowego z mechanizmem wykrywania usług itp.

Poważnym zagrożeniem dla bezpieczeństwa i wydajności systemów teleinformatycznych są ataki uzurpacji uprawnień określane tu metodą *Fałszywego VIPa*, w których Nadawca ustawia wyższą klasę żądaną niż klasa natywna obiektu, by zapewnić Klientowi nienależny poziom QoS. Powoduje to jednak nadużycie zasobów Serwera. Rozważmy np. sieć pakietową, w której węzeł źródłowy (Nadawca) umieszcza znaczniki ruchu strumieniowego (*voice /video*) w pakietach *best-effort* generowanych przez lokalną aplikację (Klienta). W następstwie tego pakiety uzyskują nienależny priorytet w kolejkach najbliższego węzła tranzytowego (Odbiorcy) i wszystkich kolejnych węzłach na trasie (Serwery). Znanym przykładem jest atak metodą podmiany klasy ruchu (ang. *traffic remapping attack*, TRA) [1] w sieciach bezprzewodowych wykorzystujących funkcję EDCA protokołu MAC IEEE 802.11 [2]. W rozproszonej bazie danych lub systemie współdzielenia plików podobny atak może przeprowadzić inteligentny terminal (Nadawcę) wystawiający w imieniu Klienta żądanie transakcji z fałszywą flagą określającą pilność lub charakter transakcji; Klient uzyskuje zatem nienależnie wysoki poziom QoS, np. dostęp do szybszego procesora lub bogatszego repozytorium zasobów w Serwerze.

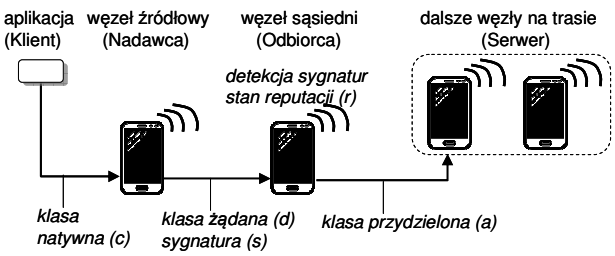
W powyższych przykładach Odbiorca przekazuje otrzymany obiekt do Serwera w niezmienionej postaci (klasa żądana pokrywa się z przydzieloną), tj. *ufa* obiektowi. W obliczu ataków Odbiorca powinien samodzielnie wnioskować o natywnej klasie obiektu, np. stosując *detekcję sygnatur*. Sygnaturą obiektu nazywamy tu zestaw cech zależnych od klasy natywnej, których Nadawca nie może modyfikować, np. długość lub format pakietu, zawartość pól danych, certyfikaty Klienta, kontekst zapytania, kod transakcji, dane dotyczące zabezpieczeń *end-to-end* itp. Detekcja sygnatur jest na ogół kosztowna: w sieciach pakietowych wymaga analizy zawartości

pakietów (ang. *deep packet inspection*, DPI), np. poprzez dopasowywanie sekwencji bitowych [3], trudnej do realizacji przy dużych prędkościach transmisji i umożliwiającej niszczące ataki typu odmowy usług. Detekcja sygnatur jest na ogół niedoskonała: pakiety ruchu strumieniowego (głos/obraz) są zwykle krótkie, ale mogą takimi być też pakiety ruchu *best-effort*; certyfikaty Klienta lub kody transakcji mogą być identyczne dla innej klasy natywnej itd.

Trudność problemu ataków metodą Fałszywego VIPa wynika z kilku oczywistych postulatów:

(i) atak jest bezkosztowy dla Nadawcy (o ile nie nakłada się płatności za samo żądanie wysokiego poziomu QoS, co jest mało praktyczne),
(ii) Odbiorca nie ma możliwości niezawodnego rozpoznania klasy natywnej obiektu: detekcja sygnatury, nawet jeśli została uruchomiona pomimo wysokiego kosztu, jest niedoskonała,
(iii) decyzja Odbiorcy odnośnie uruchomienia detekcji sygnatury nie może być uzależniona od klasy żądanej, ponieważ ta ostatnia może być przez Nadawcę zintegrowana z sygnaturą, bądź zostać ujawniona dopiero po podjęciu decyzji,
(iv) Odbiorcy nie zezwala się na oszustwa w przydziale klas, tzn. przydział niskiej klasy po wykryciu sygnatury wysokiej klasy natywnej, ponieważ mogłoby to zagrozić misji systemu,
(v) świadczenie wysokiego poziomu QoS jest kosztowne dla Serwera i na ogół nieskompensowane jakąkolwiek płatnością dla Odbiorcy,
(vi) Nadawca (i Klient) nie mają możliwości poznania klas przydzielanych przez Odbiorcę konkretnym obiektom – ich percepcja jakości usług formuje się jedynie po długiej sekwencji takich przydziałów.

W przykładzie ataku TRA (rys. 1) postulat (i) jest oczywisty – atak polega na prostej podmianie pola w nagłówku pakietu; postulat (ii) wynika stąd, że cechy niektórych pakietów ruchu *best-effort* (np. długość, adresy portów, określone sekwencje bitowe) mogą być typowe dla ruchu strumieniowego i odwrotnie; postulat (iii) nabiera znaczenia, gdy DPI wykonuje się w trybie *cut-through*, tj. w miarę odbioru kolejnych bitów pakietu, zaś klasa żądana znajduje się we fragmencie końcowym; postulat (iv) wynika ze szkód związanych z nieprioritetowym traktowaniem ruchu strumieniowego; postulat (v) jest realistyczny, gdyż priorytetowe traktowanie pakietów wykonujących atak TRA zużywa pasmo przeznaczone dla innych klientów; wreszcie postulat (vi) jest realistyczny, ponieważ węzeł źródłowy nie może odczytywać nagłówków (przydzielonych priorytetów) pakietów przekazywanych w dalszej części trasy.



Rys. 1. Model ataku na przykładzie TRA

Postulaty (i) i (ii) stwarzają zachętę dla Nadawcy do ataków metodą Fałszywego VIPa bez kosztów i ryzyka wykrycia. Postulat (iii) uniemożliwia Odbiorcy łatwe oszczędności, np. uruchamianie detekcji sygnatury tylko gdy klasa żądana jest wysoka. Z uwagi na postulat (iv) Odbiorca nie może łatwo ukarać ataku, który podejrzewa (np. gdy sygnatura nie pasuje do klasy żądanej). Postulaty (v) i (vi) tworzą dla Odbiorcy tzw. pokusę nadużycia (ang. *moral hazard* [4]), tj. oszukiwania z naruszeniem postulatu (iv). Końcowa uwaga w postulacie (vi) oznacza, że przydział klas u Odbiorcy postrzegany jest przez Nadawcę jedynie w sensie statystycznym; stąd systematyczne naruszanie postulatu (iv) może wywołać jego podejrzenia. Postulat (vi) wyklucza również zastosowanie przez Nadawcę algorytmów bieżącej predykcji dla odgadnięcia zasad przydziału klas przez Odbiorcę, a tym samym wypracowania skutecznych strategii ataku.

Przy tak skąpej informacji co do działania Odbiorcy, Nadawca może uciec się do probabilistycznej strategii ataku: atak na dany obiekt następuje z prawdopodobieństwem zależnym od jego klasy natywnej, optymalizowanym z punktu widzenia wartości oczekiwanej postrzeganego poziomu QoS. Bez niezawodnej obserwacji klasy natywnej obiektu (postulat (ii)) Odbiorcy trudno jest zidentyfikować strategię ataku Nadawcy – wnioskowanie na podstawie statystyk klas żądanych oraz wykrytych sygnatur wymagałoby znajomości statystyk generacji klas natywnych, które są prywatną informacją Nadawcy. Odbiorca powinien odpowiednio wyważać koszty świadczenia QoS oraz uruchamiania detekcji sygnatury, z zachowaniem zadowalającej statystycznej percepcji jakości usług u Nadawcy/Klienta.

Jako możliwe podejście proponuje się tutaj wprowadzenie *podsystemu reputacyjnego* u Odbiorcy. Pomaga on Odbiorcy w decyzjach o uruchomieniu detekcji sygnatury dla otrzymanego obiektu (dla zmniejszenia kosztu przydziału nienależnie wysokiego poziomu QoS), bądź zaufaniu mu (dla zmniejszenia kosztu detekcji sygnatur). Wyróżnia się pewną liczbę stanów reputacji, aktualizowanych od obiektu do obiektu, przy czym jedynie najwyższy spośród nich (tzw. *stan zaufania*) pozwala pominąć detekcję sygnatury. Porównanie klasy żądanej i sygnatury obiektu otrzymanego w innym stanie reputacji determinuje aktualizację stanu reputacji. Aby utrudnić stosowanie bardziej zaawansowanych strategii ataków, aktualny stan reputacji nie jest ujawniany Nadawcy. Podsystem reputacyjny jest więc *obustronnie ślepy*: nie może on obserwować prawdziwych zachowań Nadawcy/Klienta, który z kolei nie może obserwować stanów reputacji oraz wynikających z nich przydziałów klas.

Dalej dla uproszczenia rozróżnimy tylko dwie klasy natywne: L (niską) i H (wysoką). Łatwo wówczas zdefiniować *użyteczności* (tj. miary korzyści bądź kosztów) Nadawcy i Odbiorcy: korzyść Nadawcy odzwierciedla skuteczność ataków; koszt Odbiorcy związany jest z dostarczaniem wysokiego poziomu QoS (ponoszonego przez Serwer) oraz uruchamianiem detekcji sygnatur.

2. WCZEŚNIEJSZE PRACE

Istniejące podejścia do obrony przed TRA można w ramach naszego opisu podsumować następująco. W [5] Odbiorca obciąża Nadawcę dodatkową płatnością, gdy klasą żądaną jest H. Jednak system płatności może być trudny do wdrożenia; proponowane w nieniejszej pracy rozwiązanie nie wymaga go. W [6] Odbiorca narzuca Serwerowi bardziej równoprawne traktowanie obiektów klasy L w stosunku do klasy H. Jednakże taka semantyka może nie być wspierana przez protokół komunikacyjny Odbiorca-Serwer lub mechanizmy świadczenia usług w Serwerze. Podejście przedstawione w [7] wymaga, by przydzielona przez Odbiorcę klasa H była niezależnie zweryfikowana przez Serwer. W innym rozwiązaniu [1] Odbiorca sygnalizuje dyskomfort z powodu częstego przydziału klasy H i gotowość do uruchomienia mechanizmu penalizacji Nadawcy, np. przydziału klasy L kolejnym obiektom wbrew wykrywanym sygnaturom. W [8] Nadawca dokonuje detekcji sygnatur na rzecz Odbiorcy zamiast działać na rzecz Klienta.

W niniejszej pracy podsystem reputacji wspomaga obronę Odbiorcy przed atakami metodą Fałszywego VIPs, stwarzając mu namiastkę uczenia się zachowań Nadawcy. Odbiorca winien przy tym ukrywać zarówno zasady działania podsystemu reputacji, jak i aktualny stan reputacji. Przypomina to koncepcję ukrywania mechanizmów bezpieczeństwa (ang. *security by obscurity* [9]). Jednak Nadawca może próbować bieżącej predykcji (ang. *online prediction* [10], [11]) stanu reputacji. W modelu ogólnym jakość predykcji odpowiedzi na kolejne pytania poprawia się w oparciu o poprawne odpowiedzi na poprzednie pytania. Odpowiedni model teoriogrowy opisano w [12]: aby uzyskać wysoką użyteczność, uczący się musi dopasować swą akcję do akcji przeciwnika, odgadując ją na podstawie przeszłych akcji (Nadawca szuka koincydencji klasy żądanej L i przydzielonej H). W naszym kontekście problem z uczeniem polega na tym, że ani sygnatury, ani klasy przydzielone poszczególnym obiektom nie są obserwowalne dla Nadawcy. Teoriogrowe modele systemów wykrywania włamań (ang. *intrusion detection systems*, IDS [13]) są także mało pomocne: w związku z postulatem (iv) ryzyko ujawnienia się jako atakujący nie istnieje, zatem atak jest dla Nadawcy strategią dominującą.

3. MODEL SYSTEMU

Nadawca i Odbiorca są połączeni kanałem komunikacyjnym (por. rys. 1). Obiekty o klasie natywnej L lub H, generowane sekwencyjnie przez Klienta, przekazywane są przez Nadawcę do Odbiorcy jako żądania usług. Generacją obiektów rządzi stacjonarny bezpamięciowy proces przypadkowy. Niech $c^{(k)}$ oznacza klasę natywną k-tego generowanego obiektu, $k = 1,2,...$, i niech ρ będzie prawdopodobieństwem, że wygenerowany obiekt ma klasę natywną H. Oznaczmy przez $rand(\pi)$ zdarzenie losowe występujące z prawdopodobieństwem π. Mamy $c^{(k)} = \mathrm{H}$, gdy $rand(\rho)$ i $c^{(k)} = \mathrm{L}$, gdy $\neg rand(\rho)$.

Generowany obiekt posiada sygnaturę, tj. zbiór cech, które Odbiorca definiuje jako istotne z punktu widzenia przydziału klasy (sama klasa natywna nie jest obserwowalna dla Odbiorcy). Nadawca nie ma możliwości modyfikacji sygnatury, jest jej jednak świadom, jeśli zna zasady detekcji sygnatur u Odbiorcy. Realistyczne jest założenie, że sygnatura zależy od klasy natywnej, choć nie w sposób deterministyczny: obiekt klasy natywnej L może "przypadkowo" mieć sygnaturę charakterystyczną dla klasy H i *vice versa*. Niech $s^{(k)}$ reprezentuje sygnaturę k-tego obiektu, zaś $\varepsilon_c = \Pr[s^{(k)} \neq c^{(k)} \mid c^{(k)} = c]$ będzie stopą błędów sygnatury, tzn. prawdopodobieństwem, że generowany obiekt ma "nieprawidłową" sygnaturę. Mamy $s^{(k)} \neq c^{(k)}$, gdy $rand(\varepsilon_{c^{(k)}})$ i $s^{(k)} = c^{(k)}$ w przeciwnym razie. Przekazując k-ty obiekt Odbiorcy, Nadawca ustawia klasę żądaną $d^{(k)}$ zależnie od (znanej sobie) klasy natywnej $c^{(k)}$, sygnatury $s^{(k)}$ oraz stacjonarnej strategii $(\sigma_\mathrm{L}, \sigma_\mathrm{H})$ ataku metodą Fałszywego VIPa, gdzie $\sigma_c = \Pr[d^{(k)} = \mathrm{H} \mid s^{(k)} = \mathrm{L} \wedge c^{(k)} = c]$; naturalnie w przypadku gdy $s^{(k)} = \mathrm{H}$, jedynym sensownym ustawieniem jest $d^{(k)} = \mathrm{H}$. Mamy więc $d^{(k)} = \mathrm{H}$, gdy $s^{(k)} = \mathrm{H} \vee rand(\sigma_{c^{(k)}})$, w przeciwnym razie $d^{(k)} = \mathrm{L}$.

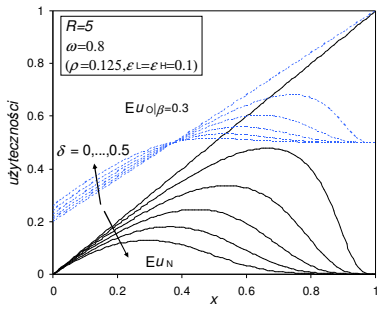
Przydział klasy (tj. poziomu QoS) dla k-tego obiektu u Odbiorcy zależy od wykrytej sygnatury oraz stanu reputacji $r^{(k)}$ Nadawcy tuż przed otrzymaniem obiektu. Niech $r^{(k)} \in \{1,\ldots,R\}$, zaś klasę przydzieloną oznaczmy przez $a^{(k)}$, gdzie $k = 1,2,\ldots$ i $R \geq 2$. Stanem zaufania jest R, kiedy to Odbiorca przekazuje obiekt do Serwera w postaci niezmienionej, tj. $a^{(k)} = d^{(k)}$; w pozostałych stanach przydział klasy zależy od wykrytej sygnatury: $a^{(k)} = d^{(k)}$, gdy $r^{(k)} = R$ oraz $a^{(k)} = s^{(k)}$, gdy $r^{(k)} < R$ (zauważmy, że niemożliwa jest sytuacja $(s^{(k)},a^{(k)}) = (\mathrm{H},\mathrm{L})$). W sytuacji gdy $r^{(k)} < R$ i $(s^{(k)},d^{(k)}) = (\mathrm{L},\mathrm{H})$, Odbiorca podejrzewa atak (nie ma pewności, gdyż $s^{(k)} = \mathrm{L}$ nie implikuje $d^{(k)} = \mathrm{L}$) i obniża stan reputacji Nadawcy. Natomiast uczciwe żądanie klasy L, tj. $(s^{(k)},d^{(k)}) = (\mathrm{L},\mathrm{L})$, podnosi stan reputacji. Gdy $r^{(k)} = R$, Odbiorca nie obserwuje $s^{(k)}$ ani $d^{(k)}$ i obniża stan reputacji z prawdopodobieństwem $\Pr[d^{(k)} = \mathrm{H}]$, które szacuje na podstawie obiektów otrzymanych w przeszłości. W innych przypadkach stan reputacji nie zmienia się. Oprócz R Odbiorca definiuje parametr $\delta \in [0, 1]$, który wyraża tendencję do obniżania stanu reputacji; R i δ stanowią prywatną informację Odbiorcy. Formalnie, $r^{(k+1)} = r^{(k)} - 1$, gdy $rand(\delta)$ oraz $(r^{(k)} = R$ i $rand(\Pr[d^{(k)} = \mathrm{H}]))$ lub $(1 < r^{(k)} < R$ i $(s^{(k)},d^{(k)}) = (\mathrm{L},\mathrm{H}))$; $r^{(k+1)} = r^{(k)} + 1$, gdy $r^{(k)} < R$ i $(s^{(k)},d^{(k)}) = (\mathrm{L},\mathrm{L})$ i $\neg rand(\delta)$; w pozostałych przypadkach $r^{(k+1)} = r^{(k)}$.

4. UŻYTECZNOŚCI

Przyjmiemy, że generacja k-tego obiektu przez Klienta przynosi Nadawcy jednostkową korzyść w razie skutecznego ataku metodą Fałszywego VIPa i jednostkową stratę w razie niesłusznie przydzielonej klasy L: $u_\mathrm{N}(c^{(k)},a^{(k)}) = +1$ gdy $(c^{(k)},a^{(k)}) = (\mathrm{L},\mathrm{H})$ i $u_\mathrm{N}(c^{(k)},a^{(k)}) = -1$ gdy $(c^{(k)},a^{(k)}) = (\mathrm{H},\mathrm{L})$; inaczej $u_\mathrm{N}(c^{(k)},a^{(k)}) = 0$. Natomiast użyteczność Odbiorcy definiujemy jako liniową kombinację kosztów detekcji sygnatur (gdy obiekt jest otrzymany w stanie różnym od R) i dostarczania wysokiego poziomu QoS (gdy klasą przydzieloną jest H), tj.

$u_O = \beta \cdot 1_{r^{(k)}<R} + 1_{a^{(k)}=H}$, gdzie $\beta > 0$ jest względną wagą kosztów detekcji sygnatur, zaś $1_Z = 1$, gdy Z = `true` oraz $1_Z = 0$, gdy Z = `false`.

Nadawca optymalizuje swą strategię (σ_L, σ_H), zaś Odbiorca swoje parametry (R, δ) z punktu widzenia oczekiwanych wartości użyteczności przy długich sekwencjach generowanych obiektów. Analiza łańcucha Markowa $(r^{(k)})_{k=0,1,2,\ldots}$ pokazuje, że względna nadwyżka oczekiwanego zysku Nadawcy w stosunku do sytuacji braku podsystemu reputacyjnego wynosi $\mathsf{E}u_N = x\pi_R = x \Big/ \left(1 + \left(\frac{1}{\omega(1-x)} - 1\right) \cdot \frac{\delta}{1-\delta} \cdot \sum_{i=0}^{R-2} \left(\frac{\delta}{1-\delta} \cdot \frac{x}{1-x}\right)^i\right) \in [0, 1]$, zaś oczekiwany koszt Odbiorcy $\mathsf{E}u_O = \beta(1 - \pi_R) + 1 - \omega(1 - \mathsf{E}u_N)$, gdzie π_R jest stacjonarnym prawdopodobieństwem stanu zaufania, $\omega = (1 - \rho)(1 - \varepsilon_L) + \rho\varepsilon_H$, natomiast strategię ataku reprezentuje $x = \big((1-\rho)(1-\varepsilon_L)\sigma_L + \rho\varepsilon_H\sigma_H\big)/\omega$. Rys. 2 przedstawia oczekiwane użyteczności w funkcji x dla $R = 5$, $\omega = 0.8$, $\beta = 0.3$ i różnych δ.



Rys. 2. Oczekiwane użyteczności Nadawcy i Odbiorcy

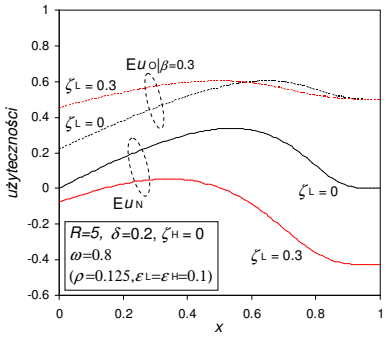
5. WPŁYW NIEPEWNOŚCI NADAWCY CO DO SYGNATUR OBIEKTÓW

W bardziej realistycznym modelu Nadawca jedynie przewiduje $s^{(k)}$ przy braku pewności. Niech dla k-tego obiektu $p^{(k)}$ będzie *przewidywaną klasą przydzieloną* przez Odbiorcę i stanowiącą dla Nadawcy podstawę ustawienia $d^{(k)}$. Niech $\varsigma_L = \Pr[p^{(k)} = H \mid s^{(k)} = L]$ i $\varsigma_H = \Pr[p^{(k)} = L \mid s^{(k)} = H]$ (poprzednio $\varsigma_L = \varsigma_H = 0$, tj. $p^{(k)} \equiv s^{(k)}$). Mamy więc $d^{(k)} = H$, gdy $p^{(k)} = H$ lub $rand(\sigma_{c^{(k)}})$, w przeciwnym przypadku $d^{(k)} = L$. Analiza markowowska pokazuje teraz, że $\mathsf{E}u_N = 1 - \omega/\Omega + (y + \omega/\Omega - 1)\pi_R$, przy czym $y = \frac{\omega}{\Omega}\Big((1-\rho)\big((1-\varepsilon_L)(1-\varsigma_L) + \frac{1-\omega}{\omega}\varepsilon_L\varsigma_H\big)\sigma_L + \rho\big(\varepsilon_H(1-\varsigma_L) + \frac{1-\omega}{\omega}(1-\varepsilon_H)\varsigma_H\big)\sigma_H\Big)$, zaś $\Omega = \omega(1 - \varsigma_L) + (1 - \omega)\varsigma_H$; prawdopodobieństwo stanu zaufania wynosi tutaj $\pi_R = 1 \Big/ \left(1 + \frac{\delta}{1-\delta}\left(\frac{1}{\Omega(1-y)} - 1\right)\sum_{i=0}^{R-2}\left(\frac{\delta}{1-\delta} \cdot \frac{\omega}{\Omega} \cdot \frac{(1-\varsigma_L)x+\varsigma_L}{\Omega(1-y)}\right)^i\right)$, gdzie ω i x mają ten sam sens co wcześniej. Oczekiwany koszt Odbiorcy wynosi $\mathsf{E}u_O = \beta(1-\pi_R) + 1 - \Omega\big(1 - \mathsf{E}u_N\big)$. Zauważmy, że w odróżnieniu od sytuacji $\varsigma_L = \varsigma_H = 0$ obie użyteczności zależą od σ_L i σ_H z osobna (poprzez x i y). Ponadto, choć Odbiorca nie narusza postulatu (iv), sytuacja $(s^{(k)}, a^{(k)}) = (H,L)$ jest teraz możliwa; jej prawdopodobieństwo jest nie większe niż $(1 - \omega)\varsigma_H$.

Rys. 3 pokazuje wpływ $\varsigma_L > 0$ na oczekiwane użyteczności Nadawcy i Odbiorcy przy $\varsigma_H = 0$. Interesujące jest że wpływ ten jest generalnie niekorzystny dla Nadawcy, lecz nie musi polepszać kosztu Odbiorcy.

6. WNIOSKI

W referacie dokonano formalizacji ataków metodą Fałszywego VIPa w systemach teleinformatycznych wspierających różnicowanie poziomu QoS oraz zaproponowano mechanizm obronny w postaci obustronnie ślepego podsystemu reputacyjnego. Wstępna analiza Rys. 2 i 3 ujawnia interesującą możliwość uodporniania Odbiorcy na przyjętą przez Nadawcę strategię ataku, gdyż odpowiedni dobór R i δ może zmniejszać koszt Odbiorcy przy optymalnej strategii (x) Nadawcy. Planowane zastosowanie modelu tzw. gry Stackelberga [4], prawdopodobnie przyniesie bardziej precyzyjne wnioski.



Rys. 3. Wpływ niepewności Nadawcy co do sygnatur generowanych obiektów

PODZIĘKOWANIE

Praca częściowo sfinansowana przez Narodowe Centrum Nauki (umowa UMO-2016/21/B/ST6/03146).

LITERATURA

[1] Konorski J., S. Szott. 2014. „Discouraging traffic remapping attacks in local ad hoc networks". *IEEE Trans. on Wireless Comm.*, 13 (7): 3752–3767.

[2] Mangold S. *et al*. 2003. „Analysis of IEEE 802.11e for QoS support in Wireless LANs". *IEEE Wireless Comm.*, 10 (6): 40–50.

[3] Po-Ching Lin *et al*. 2008. „Using string matching for deep packet inspection". *Computer*, 41 (4): 23–28.

[4] Rasmussen E. 2001. *Games and Information*, 3rd ed., Blackwell Publishers.

[5] Cheung M. H. *et al*. 2009. „Random access protocols for WLANs based on mechanism design". *Proc. IEEE ICC'09*, Dresden, Germany.

[6] Nguyen S. H., L. L. Andrew, H. L. Vu. 2011. "Service differentiation without prioritization in IEEE 802.11 WLANs". *Proc. 36th IEEE Conf. on Local Computer Networks*, Bonn, Germany.

[7] Galluccio L. 2009. „A game-theoretic approach to prioritized transmission in wireless CSMA/CA networks". *Proc. 69th IEEE VTC*, Barcelona, Spain.

[8] Li M., B. Prabhakaran. 2005. „MAC layer admission control and priority re-allocation for handling QoS guarantees in non-cooperative wireless LANs".

Springer Mobile Networks and Applications, 10 (6): 947–959.
[9] Kerr R., R. Cohen. 2009. „Smart cheaters do prosper: Defeating trust and reputation systems". *Proc. AAMAS'09*, Budapest, Hungary.
[10] Bartlett P. L. 2015. *Online Prediction*. stat.berkeley.edu/~bartlett/papers/b-ol-16.pdf.
[11] Shalev-Shwartz S. 2007. *Online Learning: Theory, Algorithms, and Applications*. *Ph. D. thesis,* The Hebrew University of Jerusalem.
[12] Freund Y. *et al*. 1995. „Efficient algorithms for learning to play repeated games against computationally bounded adversaries". *Proc. 36th Symp. Foundations of Computer Science*, Milwaukee, WI.
[13] Patcha A., Jung-Min Park. 2006. „A game theoretic formulation for intrusion detection in MANets". *Int. J. of Network Security*, 2 (2): 131–137.