Symulacja w Badaniach i Rozwoju
Vol. 1, No. 2/2010

Rafał KASPRZYK
Wojskowa Akademia Techniczna, Wydział Cybernetyki
E-mail: rkasprzyk@wat.edu.pl

Symulator rozprzestrzeniania się złośliwego oprogramowania w sieciach komputerowych

1 Wprowadzenie

W otaczającej rzeczywistości sieci rozumiane jako zbiory wierzchołków i gałęzi reprezentujących relacje pomiędzy wierzchołkami są wszechobecne. Większość sieci rzeczywistych najczęściej powstających spontanicznie ze względu na liczbę występujących w nich węzłów i powiązań pomiędzy nimi nazywane są sieciami złożonymi (ang. *Complex Network*), rozległymi lub sieciami wielkiej skali. Od lat 50-tych minionego stulecia, sieci złożone, bez widocznej zasady organizacyjnej, były opisywane przy wykorzystaniu grafów losowych (ang. *random graphs*) [1]. Powstanie pod koniec XX wieku ogólnodostępnych baz danych gromadzących informacje o topologiach sieci rzeczywistych ujawniło jednak, że mają one szereg specyficznych cech, które nie były do tej pory uwzględniane. Okazało się więc, że chodź „otaczają nas sieci", które od wielu lat są badane i opisywane to nadal nie znamy ich topologii, a tym bardziej nie rozumiemy zasad rządzących ich dynamiką, czy ewolucją. Analizy prowadzone na rzeczywistych sieciach wykazały istnienie właściwości, które nie dają się modelować za pomocą grafów losowych. W szczególności mówi się o takich cechach sieci złożonych jak względnie niewielka liczba krawędzi (graf rzadki), relatywnie mała średnica, zaskakująco mała średnia odległość pomiędzy wierzchołkami (hipoteza „*six degree of seperation*"), wysoki stopień klasteryzacji, czy w końcu potęgowy rozkład stopnia wierzchołka [9]. Linie elektryczne, połączenia komunikacyjne, sieci internetowe, WWW, sieci społeczne, ale też połączenia neuronów w mózgu, czy cykle pokarmowe (kto kogo zjada w ekosystemie) należą do kategorii sieci złożonych. Wydaje się, że istnieje specjalny rodzaj architektury grafu specyficzny dla większości naturalnych sieci. Badania tej architektury są w toku, ale już widać, że taki model opisuje zadziwiająco wiele, często pozornie niezwiązanych ze sobą, aspektów rzeczywistości.

Pod koniec XX wieku wielu badaczy zainteresowało się szczególnie siecią Internet ze względu na jej zadziwiająco szybki wzrost dostępność i istotność dla społeczeństwa. Zadziwiająca była dla nich topologia Internetu, którego przecież rozwój nie został poddany żadnym regulacjom czy planom architektonicznym. Interesująca okazuje się również ewolucja tej globalnej sieci wymiany danych. Rozwój Internetu można określić jako spontaniczny i niekoordynowany. Niektórzy badacze odnajdują wręcz podobieństwa Internetu do żywego organizmu, a jego rozwój określają jako organiczny [12].

2 Generatory sieci

Kluczową rolę w symulowania rozprzestrzeniania się złośliwego oprogramowania (ang. *malwares*) odgrywają dobre modele i generatory sieci komputerowych. Przez dobre rozumiane są tu generatory, które kreują sieci złożone, a więc posiadające wcześniej wspomniane cechy właściwe dla sieci rzeczywistych. W literaturze wyróżnia się dwa kluczowe modele zaproponowane pod koniec XX wieku, które znane są jako sieci małego świata (ang. *Small World*) oraz sieci bezskalowe (ang. *Scale Free*) oraz liczne ich modyfikacje, rozszerzenia i uogólnienia.

Pojęcie sieci typu *Small World* wyłania się w naturalny sposób jako bardzo realistyczny, niejako pośredni przypadek pomiędzy dwoma skrajnymi sieciami: siecią regularną (ang. *Regular Network*), zwaną niekiedy pierścieniową (ang. *Ring Network*) oraz siecią losową (ang. *Random Network*). Zaczynając rozważania od sieci regularnej, okazuje się że każdy jej węzeł jest powiązany z tą samą liczbą węzłów „sąsiadów". Powstaje pytanie, czy tego typu, idealnie regularne sieci są zjawiskiem częstym. Przecież nawet w krystalografii zdarzają się „defekty", zaburzające monotonną strukturę powiązań. Praktyka dowodzi, że sieci regularne są oczywiście idealizacją rzeczywistości, co znacznie redukuje ich przydatność w tworzeniu modeli sieci rzeczywiście występujących w naturze i cywilizacji. Z drugiej strony sieć losowa wykazuje całkowity brak regularności w liczbie powiązań między węzłami, jak również brak jakiegokolwiek strukturalnego uporządkowania tych powiązań. W obu przypadkach nazwy sieci dość wiernie oddają ich naturę – pełna regularność i całkowita przypadkowość. Watts i Strogatz [2] zauważyli, że dokonując pewnego zabiegu na sieci regularnej można otrzymać modele sieci, które spotykane są w rzeczywistych systemach. Nie są one bowiem ani doskonale regularne, ani zupełnie losowe, a można je budować poprzez zastosowanie tzw. „przepinania" (ang. *rewiring*) niektórych gałęzi sieci. Modele sieci *Small World* przyczyniły się do rozpowszechnienia hipotezy „*six degree of seperation*", według której każde dwie osoby na świecie są połączone drogą zawierającą sześć społecznych powiązań. Pokazuje się również dlaczego ta najkrótsza droga jest spontanicznie odnajdywana[3].

Sieci *Scale Free* [4] można scharakteryzować, używając żargonu z obszaru sieci komputerowych jako sieci powiązań, w której kluczową rolę pełnią huby. Tego typu węzły określane są jako „*super-spreaders*". Barabasi i Albert dostrzegli, że sieci takie „rosną przez dodawanie" kolejnych węzłów według określonej hierarchii. Prawdopodobieństwo wystąpienia połączenia pomiędzy nowym węzłem, a każdym węzłem należącym do sieci bardzo silnie zależy od posiadanej już przez te węzły liczby krawędzi k i wynosi $P(k) \sim k^{-\gamma}$. Wykładnik potęgi, a więc γ, zależy od rodzaju rozpatrywanej sieci. Jako następstwo powyższego, sieci typu *Scale Free* „rosną przez dodawanie" kolejnych węzłów według określonej hierarchii. Tak więc, najwięcej nowych węzłów łączy się z węzłami, które posiadają już najwięcej sąsiadów (połączeń z innymi węzłami). Cecha ta jest określana jako „dołączenia preferencyjne" (ang, *preferential attachment*). Istnieje wiele modyfikacji algorytmów generujących sieci *Scale Free* [9]. Modyfikacje polegają głównie na zmianie możliwości ewolucji sieci, i tak np. w kolejnych krokach ewolucji możemy mieć do czynienia nie tylko z dodaniem nowego węzła wraz z nowymi krawędziami, ale również z dodaniem jedynie krawędzi do już istniejących węzłów, czy z przepięciem wybranych krawędzi.

3 Miary centrów grafu

Analiza dynamiki rozprzestrzeniania się złośliwego oprogramowania w sieciach komputerowych i próba przeciwdziałania epidemii nie jest możliwa bez zdefiniowania miar centralności (istotności) węzłów [8]. Wprowadzone miary istotności węzłów, dzięki istnieniu wyraźnej interpretacji fizycznej, pozwalają na ciekawą analizę topologii sieci teleinformatycznych tj. precyzyjne klasyfikowanie istniejących/projektowanych sieci, ze względu na ich niezawodność, z uwzględnieniem kontekstu stanowiącego cel istnienia/projektowania sieci. Wierzchołki centralne sieci są szczególnie interesujące ponieważ pełnią one kluczowe role stanowiąc swego rodzaju katalizatory epidemii tj. węzły, które wpływają znacząco na jej dynamikę tzw. „*super-spreaders*". Miary centralności ułatwiają udzielenie odpowiedzi na pytanie „kto (co) jest najważniejszy(e) w analizowanej sieci?". Okazuje się, że nie ma jednoznacznej odpowiedzi. Wszystko zależy od przyjętej semantyki słowa „istotny", co obrazuje pięć wprowadzonych dalej miar centralności.

Stopień wierzchołka (and. *degree centrality*)

Najprostszą miarą centralności wierzchołka, którą podpowiada intuicja, jest stopień wierzchołka. Według tej miar wierzchołek jest tym istotniejszy dla sieci im ma więcej bezpośrednich połączeń z pozostałymi wierzchołkami w sieci.

$$center_i^{Degree} = \frac{k_i}{n-1} \tag{1}$$

gdzie k_i oznacza stopień i-tego węzła, n – liczba wierzchołków sieci.

Promień wierzchołka (ang. *radius centrality*)

Jeśli wierzchołek jest tym ważniejszy im jego odległość do najdalszego wierzchołka jest najmniejsza, wówczas należy zastosować miarę istotności wyliczaną w oparciu o promień wierzchołka. Algorytm odnalezienia węzła centralnego składa się z dwóch kroków. W pierwszej pętli dla każdego wierzchołka w sieci wyliczamy jego promień, a następnie w drugim kroku szukamy węzła dla którego promień jest najmniejszy.

$$center_i^{Radius} = \frac{1}{max\, d_{ij}} \tag{2}$$

gdzie d_{ij} najkrótsza droga wyrażana minimalną liczbą krawędzi, które łączą v_i z v_j.

Średnia odległość wierzchołka (ang. *closeness centrality*)

Bardzo często zależy nam na tym, aby dokonany wybór w większości możliwych scenariuszy (średnio rzecz biorąc) był najlepszy. Przy takim założeniu „dobrą" miarą oceny istotności wierzchołka będzie wyliczenie jego średniej odległości do wszystkich pozostałych węzłów w sieci. Wierzchołek, który „średnio" jest najbliższy wszystkim wierzchołkom w sieci jest wówczas najistotniejszy.

$$center_i^{Closness} = \left[\frac{\sum_{j=1}^{n} d_{ij}}{n-1}\right]^{-1} = \frac{n-1}{\sum_{j=1}^{n} d_{ij}} \tag{3}$$

Obciążenie wierzchołka (ang. *betweenness/load centrality*)

Ciekawym sposobem oceny istotności wierzchołka w sieci jest wyznaczenie tzw. obciążenia węzła, które można zdefiniować jako procent najkrótszych dróg pomiędzy dowolną parą wierzchołków przechodzących przez rozpatrywany węzeł. Jeśli przez $p_{lk}(i)$ oznaczymy liczbę najkrótszych dróg pomiędzy węzłem v_l i v_k przechodzących przez węzeł v_i oraz przez p_{lk} liczbę wszystkich dróg pomiędzy węzłem v_l i v_k, wówczas:

$$center_i^{Betweenness} = \frac{2\sum_{l<k}\frac{p_{lk}(i)}{p_{lk}}}{(n-2)(n-1)} \tag{4}$$

Obciążenie wierzchołka jest niezwykle istotne dla odpowiedzi na pytanie: jak trudne i czasochłonne może być zadanie polegająca na maksymalizacji rozpojenia sieci w celu zminimalizowania możliwego obszaru objętego epidemią. Usunięcie wierzchołka o największej wartości obciążenia powoduje znaczące zwiększenie średniej odległości pomiędzy węzłami, a tym samym największe utrudnienia komunikacji w sieci.

Istotność sąsiedztwa wierzchołka (ang. *eigenvector centrality*)

O ile stopień wierzchołka za wartość istotności przyjmuje liczbę jego najbliższych sąsiadów, to w rzeczywistości oczywistym jest, że nie wszyscy sąsiedzi wierzchołka powinni zwiększać jego istotność o tą samą wartość. Jeśli wierzchołek ma znaczną liczbę połączeń, ale z wierzchołkami, które nie są istotne dla sieci to nie powinien on być uznany za tak ważny jak wierzchołek, który posiada choćby jedno połączenia, ale za to z najważniejszym węzłem w sieci. Tak więc istotność węzła v_i oznaczona jako e_i powinna zależeć od istotności wierzchołków z jakimi jest on połączony i możemy ją wyliczyć następująco:

$$e_i = \frac{1}{\lambda}\sum_{j=1}^{n} A_{ij}e_j \Rightarrow \vec{e} = \frac{1}{\lambda}A\vec{e} \tag{5}$$

Wartość λ (wartość własną macierzy) wyliczamy z $\det(A-\lambda I)=0$. Wektor $\vec{e}$ jest wektorem własnym macierzy A dla największej wartości λ.

4 Efektywność komunikacji w sieci

Do oceny stopnia uszkodzenia sieci w przypadku awarii będącej następstwem ataku bądź zdarzenia losowego, jak i efektywności strategii dystrybucji oprogramowania antywirusowego użyto miary zwanej *global connection efficiency* (GCE) [7]. Przy założeniu, że wydajność połączenia pomiędzy węzłem v_i i v_j jest odwrotnie proporcjonalna do najkrótszej drogi pomiędzy tymi węzłami mamy:

$$connection_{ij}^{efficiency} = \frac{1}{d_{ij}} \tag{6}$$

Interesująca jest efektywność komunikacji nie tyle pomiędzy wybranymi węzłami, ale w całej sieci przed i po ataku, jak i efektywności strategii szczepień węzłów sieci teleinformatycznej wg zadanego kryterium. Należy wyliczyć więc miarę *global connection efficiency* zdefiniowaną jako średnia wartość miara *connection efficiency* pomiędzy każdą parą węzłów.

$$global\_connection^{efficiency} = \frac{2\sum_{i<j} connection_{ij}^{efficiency}}{n(n-1)} = \frac{2}{n(n-1)}\sum_{i<j}\frac{1}{d_{ij}} \quad (7)$$

Warto w tym miejscy podkreślić, że o ile spadek wartości GCE w przypadku awarii pewnego węzła świadczy o zmniejszeniu efektywności komunikacji w sieci, to w przypadku porównywania efektywności strategii szczepień oznacza zmniejszenie możliwości rozprzestrzeniania się danego wirusa bez wpływu na funkcjonalność sieci jako całości. Tak więc miara GCE została wykorzystana w celu identyfikacji węzłów, które należy szczególnie chronić przed atakiem, jak również węzłów które muszą podlegać pełnej ochronie antywirusowej, gdyż mogą stać się szczególnymi katalizatorami epidemii ze względu na ich wpływ na globalną efektywność komunikacji w sieci.

5 Modelowanie rozprzestrzeniania się złośliwego oprogramowania

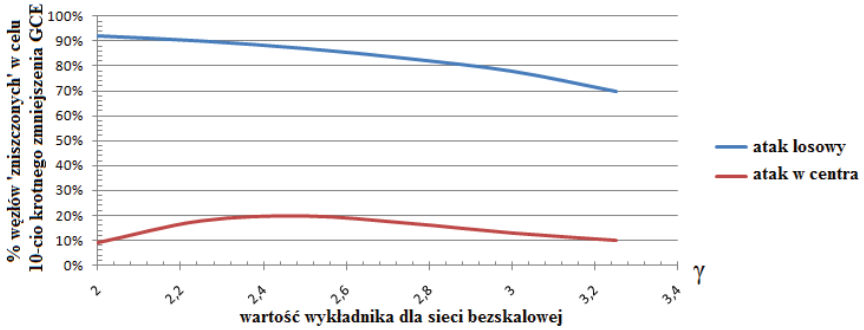
Internet, który zrodził się z chęci stworzenia solidnego systemu komunikacji mogącego przetrwać obcą agresję jest coraz częściej postrzegany jako infrastruktura niezwykle wrażliwa, od której zależy bezpieczeństwo państwa. Aby przeprowadzić skuteczny atak na tą infrastrukturę nie trzeba mobilizować sił militarnych. Człowiek wyposażony w standardowe technologie komputerowe, posiadający odpowiednią wiedzę, może przeprowadzić cyberatak o skutkach wręcz katastrofalnych dla współczesnego systemu polityczno-gospodarczego.

Jednym z zagrożeń jest przejęcie kontroli nad siecią teleinformatyczną, co wydaje się być możliwe, ale dość skomplikowane. Znacznie prostsze i prawdopodobnie równie niekorzystne byłoby przeprowadzenie mniej wyrafinowanego ataku polegającego na sparaliżowaniu sieci. Kamieniem milowym w badaniach nad odpornością Internetu na różnego rodzaju ataki było odkrycie bezskalowej topologii tej sieci. Wprowadzone miary centralności oraz współczynnik GCE pozwoliły na porównanie odporności sieci bezskalowej na awarie spowodowane losowym jak i celowym atakiem.

Przeprowadzona analiza dowodzi, że charakterystyczna długość ścieżki praktycznie wcale nie zmienia się, gdy 10% jej węzłów zostanie losowo zniszczonych, co więcej dopiero losowe zaatakowanie 80% węzłów powoduje 10-krotny spadek wartości GCE. Ponieważ większość węzłów w sieci bezskalowej ma tylko jedno lub dwa połączenia, losowe wyeliminowanie węzła nie ma większego wpływu na komunikację pomiędzy pozostałymi węzłami. Internet ma więc strukturę, jaka jest potrzeba, by zapewnić niezawodną komunikację, nawet wówczas, gdy nie wszystkie węzły działają. Lokalna niewydolność nie wynika jedynie z awarii węzłów, gdyż często węzeł może stać się tymczasowo bezużyteczny, bo jest zablokowany przez masę przechodzącej przez niego informacji. W sieci bezskalowej można łatwo znaleźć alternatywną drogę, nawet jeśli wiele węzłów jest jednocześnie zablokowanych. Co ciekawe w dowolnym momencie około 3 procent routerów w sieci Internet jest zablokowana, a sieć mimo to doskonale działa. Bezskalowa struktura Internetu gwarantuje, że próby eliminowania losowo wybranych węzłów będą prawdopodobnie nadaremne, ponieważ sieć może wytrzymać sporą liczbę takich ataków bez większej utraty zdolności komunikacji.

Mimo swej solidności i efektywności, Internet ma pewną wadę, która może okazać się fatalna w skutkach. Czy cyberterrorysta będzie bowiem próbował tak mało

wyrafinowanego ataku? O ile w sieci losowej, żaden z węzłów nie ma szczególnego charakteru. W sieci bezskalowej pewne węzły są zdecydowanie ważniejsze (węzły centralne), co powoduje, że po ich usunięciu sieć ulega dezintegracji. Okazuje się, że już usunięcie 15% węzłów centralnych powoduje 10-krotny spadek wartości GCE, a tym samym brak możliwości efektywnej komunikacji w sieci. Słabością sieci bezskalowych jest również duże prawdopodobieństwo pojawienia się kaskady awarii, gdyż usterka w jednym z centrów, przenosi obciążenie na inne węzły mogąc powodować ich przeciążenie. Zatem, aby sparaliżować Internet należy zidentyfikować stosunkowo niewielką liczbę węzłów, a następnie uczynić je celem ataku. Wniosek jest prosty, najbardziej istotne węzły należy wyjątkowo dobrze chronić tj. zbudować dobre mechanizmy ochronne i/lub należałoby uczynić je „ukrytymi".



Rys. 1.Wyniki symulacji ataku na sieć bezskalową ocenane z wykorzystaniem GCE
Fig. 1.Simulation's outcomes assess using GCE measure

Rozmyślne niszczenie węzłów, jak i ich przypadkowe awarie to nie jedyne zagrożenie dla sieciowych infrastruktur krytycznych. Wirusy komputerowe, które rzadko uszkadzają fizycznie węzły sieci, wykorzystują raczej topologie sieci, by nią „zawładnąć" dla własnych celów. Wirusy komputerowe zwykle rozprzestrzeniają się poprzez zainfekowane pakiety danych przekazywane z jednego komputera do drugiego, podobnie zresztą jak wirusy biologiczne przenoszą się pomiędzy ludźmi jako cząsteczki we wdychanym powietrzu lub przez płyny ustrojowe. Gdy pomiędzy węzłami nastąpi tak rozumiane połączenie to mogą się one od siebie zarazić. Jak się okazuje sieci teleinformatyczne, nie poddają się uodpornieniu przeciwko złośliwemu oprogramowaniu, w wyniku strategii losowego szczepienia węzłów. Bardzo szybko jednak stają się odporne w wyniku szczepienia węzłów centralnych tj. posiadających najwyższe wartości opisanych miar centralności. Obserwacja ta okazuje się niezwykle przydatna przy planowaniu szczepień mających zapobiec rozprzestrzenianiu się epidemii złośliwego oprogramowania.

Standardowym podejściem w epidemiologii jest uproszczające założenie, że choroba zakaźna rozprzestrzenia się w populacji modelowanej jako graf losowy lub regularny. Model ten przewiduje pewien próg epidemii. Choroba szerzy się w populacji, bezustannie zarażając pewien stały odsetek ludzi, jeśli tempo jej rozprzestrzeniania się jest większe od pewnej wartości progowej, a w innym wypadku szybko zanika. Wydaje

się, że niektóre epidemie faktycznie zachowują się w ten sposób. Próg epidemii ma kluczowe znaczenie, jeśli bowiem pewien odsetek populacji będzie zaszczepiony, to tempo szerzenia się epidemii pozostanie poniżej wartości progowej, a w związku z tym choroba nie przerodzi się w epidemię.

Internet nie jest jednak grafem losowym. Stefan Bornholdt i jego współpracownicy z Uniwersytetu w Kiel pokazali, że połączenia e-mailowe również tworzą sieć bezskalową, co sugeruje, że „sieć znajomych" zdefiniowana na sieci elektronicznej ma taki sam charakter jak Internet [10]. Fizycy Romualdo Pastor-Satorras w Barcelonie i Alessandro Vespignani w Trieście odkryli, że fakt bezskalowej topologii sieci po której wędrują e-maile całkowicie zmienia sposób rozprzestrzeniania się wirusów komputerowych. Korzystając z symulacji komputerowej w celu zbadania, jak zachowuje się choroba zakaźna w sieci bezskalowej, odkryli, że nie występuje tam próg szerzenia się epidemii [6]. Niezależnie jak wolne jest tempo rozprzestrzeniania się wirusa, może on przeniknąć cały system, zarażając pewien odsetek węzłów. Ponieważ zarażone węzły mogą zostać „wyleczone" dzięki programom komputerowym, wirus w końcu zanika. Proces ten zachodzi jednak bardzo powoli. Oprogramowanie zwalczające dany wirus z reguły staje się dostępny już po kilku dniach lub tygodniach od infekcji, a jednak wirusy mogą przetrwać w sieci nawet przez wiele lat.

Czy sytuacja jest tak fatalna? Jak wiadomo, powodem takiego zachowania sieci bezskalowych jest nieproporcjonalny wpływ pewnych węzłów na rozprzestrzenianie się chorób. Jeśli przerwie się połączenia do tych kluczowych węzłów to cała sieć szybko rozpadnie się. Pastor-Sattoras i Vespignami pokazali, że nakierowanie szczepień ochronnych na jednostki prowadzące bogate życie seksualne zdecydowanie obniża wrażliwość sieci na epidemie chorób przenoszonych drogą płciową [11]. Na tej samej zasadzie wybuch epidemii wirusa komputerowego może zostać skutecznie powstrzymany poprzez „zaszczepienie" zaledwie 15% wierzchołków wybranych ze względu na liczbę połączeń.

Niezwykle interesujące jest to, że Internet rozrósł się bez żadnego planu w tę, jak się wydaje, najbardziej solidną z możliwych struktur sieci. Gdyby istniała organizacja pod dyktando której należałoby dokonywać kolejnych przyłączeń węzłów to zapewne powstała struktura nie byłaby tak solidna. Czasami więc najlepiej jest pozwolić by system sam się organizował. Pozostaje jednak pytanie, dlaczego Internet ma taka właśnie strukturę?

6 Funkcjonalność aplikacji do badania odporności sieci

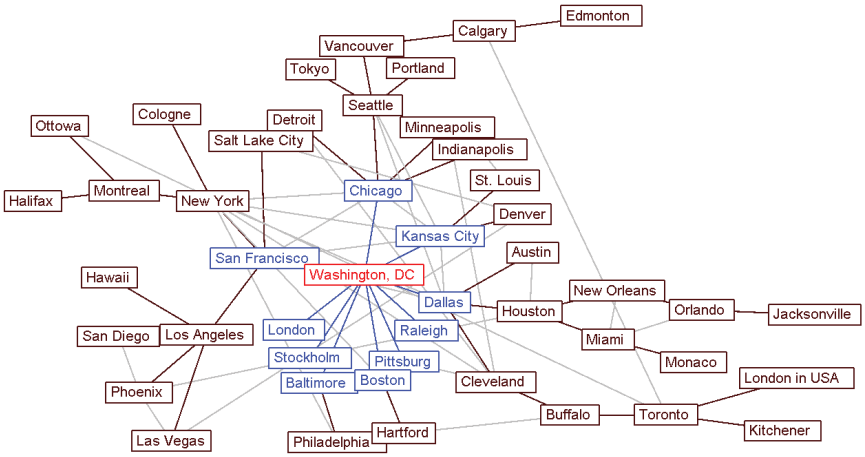
Symulator został zaimplementowany na platformie .NET 3.5 z SP1. Pierwotnie projektowany był jako aplikacja desktopowa. Ostatnie pozytywne doświadczenia z aplikacjami webowymi doprowadziły do całkowitej modyfikacji architektury. Obecnie aplikacja działa zgodnie z koncepcją SOA (ang. *Service Oriented Architecture*). Interfejs użytkownika natomiast został stworzony w oparciu o rozwiązanie AJAX (ang. *Asynchronous JavaScript and XML*) z wykorzystaniem *Microsoft Silverlight* 3.0.

Funkcjonalność aplikacji można sprowadzić do kilku kluczowych obszarów:

a) Modelowanie złośliwego oprogramowania (wirusy, robaki, trojany, itp.) mającego szkodliwe, przestępcze lub złośliwe działanie w stosunku do użytkownika komputera – moduł Mawares Models;
b) Generowanie modeli sieci o topologiach rzeczywistych sieci teleinformatycznych z wykorzystaniem zaimplementowanych generatorów sieci złożonych – moduł *Critical Infrastructure Networks*;
c) Konstrukcja kwestionariuszy wykorzystywanych do zbierania danych dotyczących topologii wirtualnej sieci kontaktów pomiędzy użytkownikami komputerów – moduł *Polls*;
d) Identyfikacja węzłów sieci teleinformatycznej istotnych z punktu widzenia odporności sieci na awarie, jak i strategii dystrybucji oprogramowania antywirusowego z wykorzystaniem zaimplementowanych miar centralności. Weryfikacja skuteczności proponowanych strategii z wykorzystaniem proponowanej miary GCE – moduł *Vaccination Strategies*;
e) Symulacja i wizualizacja możliwych scenariuszy rozprzestrzeniania się złośliwego oprogramowania modelowanego z wykorzystaniem modelu stworzonego w punkcie a) na sieci wygenerowanej w punkcie b) lub wirtualnej sieci kontaktów zidentyfikowanej za pomocą kwestionariuszy z punktu c). Scenariusze tworzone są z pomocą modułu d) – moduł *Spread Visualization*;
f) Ocena oczekiwanych skutków awarii węzłów i/lub epidemii złośliwego oprogramowania oraz szacowanie środków (programów antywirusowych) niezbędnych do przeciwdziałania lub zahamowania epidemii – moduł *Reports*.

Rys. 2. Główne okno aplikacji do badania odporności sieci teleinformatycznych
Fig. 2. Main window of the application to study network resistance

Jednym z bardziej efektownych elementów prezentowanej aplikacji jest moduł e), który prócz symulacji służy do interaktywnej wizualizacji powiązań pomiędzy węzłami sieci. Analiza tego typu pozwala na badanie topologii sieci komputerowej. Na tej podstawie ułatwione staje się zadanie identyfikacji kluczowych węzłów, analiza możliwych sposobów komunikowania się w sieci i wiele innych. Diagramy powiązań mogą służyć również do przedstawiania i analizy dużych zbiorów danych, dotyczących np. rozmów telefonicznych, transakcji na kontach bankowych czy ruchu internetowego. W takich przypadkach diagramy posłużyć mogą do łatwej identyfikacji wspólnych elementów, grup elementów ściśle powiązanych ze sobą czy też powiązań pośrednich. Waga wizualizacji powiązań wydaje się być, w tym przypadku, nie do przecenienia.



Rys. 3. Zobrazowanie sieci UUNET z wykorzystaniem zbudowanej aplikacji

Fig. 3. UUNET projection using introduced application

7 Podsumowanie

W artykule opisano wpływ topologii sieci Internet na jej wrażliwość na awarie oraz sposób rozprzestrzeniania się złośliwego oprogramowania. Zaproponowano model symulacyjny pozwalający symulować rozprzestrzenianie się złośliwego oprogramowania w sieciach o różnych topologiach, w tym sieciach złożonych. Otrzymane wyniki pozwalają odpowiedzieć na pytanie, czy i jeśli tak to kiedy, występują różnice w skutkach celowego ataku na sieć w porównaniu ze skutkami uszkodzeń spowodowanych zjawiskami losowymi. Pokazano dlaczego standardowa wiedza epidemiologiczna jest niewystarczająca do walki z wirusami komputerowymi. Przeprowadzone symulacje dowodzą również konieczności weryfikacji istniejącego wyobrażenia o szerzeniu się epidemii w sieciach społecznych.

W chwili obecnej symulator jest prototypem systemu nad rozwojem którego wciąż trwają intensywnej prace. Jednak już na obecnym etapie symulator wykazuje bogatą funkcjonalność i dowodzi możliwości praktycznego jego wykorzystania.

Literatura

1. Erdös P., Rényi A.: *On random graphs*, Publ. Math Debrecen 6 (1959), 290-297
2. Watts Duncan J., Strogatz Steven H.: *Collective dynamics of „small-world" networks*, Nature, 393:440-442, 1998
3. Kleinberg Jon M.: *Navigation in small world*, Nature 406, 845 (24 August 2000)
4. László B.A., Réka A.: *Emergency of Scaling in Random Networks*, Science, 286:509-512, 1999
5. Brandes U., Kenis P., Raab J.: *Explanation Through Network Visualization*, Methodology 2006, Vol. 2(1): 16-23.
6. Pastor-Satorras R., Vespignani A.: *Epidemic Spreading in Scale-Free Networks*, PRL Volume 86, Number 14 p. 3200 (2 April 2001)
7. Crucitti P., Latora V., Marchiori M., Rapisarda A.: *Error and attack tolerance of complex networks*, Physica A 340 (2004), 388-394
8. Wuchty S., Stadler P.F.: *Centers of complex networks*, Journal of Theoretical Biology 222 (2003), 45-53
9. László B.A., Réka A., *Topology of Evolving Networks*: Local Events and Universality, PRL Volume 85, Number 24 p.5234 (11 December 2000)
10. Ebel H., Mielsch L.I., Bornholdt S.: *Scale-free topology of email networks* Phys. Rev. E 66, 035103(R) (2002)
11. Pastor-Satorras R., Vespignani A.: *Immunization of complex networks*, PRL Volume 65, 036104 (2002)
12. Ball P.: *Critical mass: How one thing leads to another*, New York, Farrar, Straus & Giroux, 2004

Streszczenie

Wyraźny wzrost zainteresowania systemami dających się modelować z wykorzystaniem teorii grafów i sieci jest spowodowany rosnącym znaczeniem rzeczywistych sieci wielkiej skali. Badania niezawodności i odporności tych systemów na przypadkowe, jak i celowe ataki oraz trudne do przewidzenia awarie mają oczywiste znaczenie praktyczne. W artykule przedstawiono koncepcję modelowania i symulacji zagrożeń dla sieci teleinformatycznych. Zaprezentowano aplikację umożliwiającą symulację rozprzestrzeniania się złośliwego oprogramowania, badanie struktury i prognozowanie możliwych kierunków ewolucji sieci teleinformatycznych, optymalizację sposobów wykorzystania zasobów czy w końcu formułowanie możliwych procedur postępowania w sytuacjach kryzysowych np. przypadkowe awarie pewnych węzłów lub celowe ataki terrorystyczne.

The simulator of malwares spreading in telecommunication networks

Summary

The paper focuses special attention on research of Complex Networks. Complex Networks have Scale Free and Small Word features, what make them accurate model

of many networks such as telecommunication networks. These features, which appear to be very efficient for communication, favor at the same time the spreading of malwares. Based on defined centrality measures, we show how to discover the critical elements of any network.The identification of the critical elements should be the first concern in order to reduce the consequence of epidemics. We define dynamic model for the spreading of infections on networks and build application to simulate and analyse many epidemic scenarios.