УДК 658.511
DOI: 10.31673/2409-7292.2019.040611

В.А. Савченко, О.Й. Мацько, С.В. Легомінова, І.С. Полторак, В.В. Марченко

МОДЕЛЮВАННЯ КІБЕРАТАК ЗАСОБАМИ ТЕОРІЇ ГРАФІВ

У статті розглядається комплексна модель кібератаки на основі теорії графів, яка поєднує класичні уявлення щодо моделювання складних атак з розширеннями, що враховують залежності уразливостей окремих компонентів системи та мережевий статус компонентів. Наведено приклад оцінювання сценарію атаки та зроблено висновки щодо можливості застосування моделі для прогнозування наслідків атаки.

Ключові слова: кібератака, уразливість, граф атаки, граф залежності, ігрова модель, топологія системи

Вступ

Постійно зростаюча складність обчислювальних систем, додатків та розподілених баз даних зробили їх уразливими щодо складних багатоступеневих атак (зокрема APT-вторгнень). Проблема ускладнюється тим, що дуже часто інформаційні продукти розробляються відокремлено, без єдиного задуму, нашаровуючись та сполучаючись у подальшому один з одним. При цьому питання безпеки при розробці та експлуатації не є першочерговими. Уразливості інформаційних систем мають як причинні (умови проникнення у систему) так і наслідкові (результат проникнення) обставини і тому дуже часто, при сполученні окремих систем у єдине середовище, наслідкові обставини одного модуля системи є причинними для іншого.

Постановка проблеми

Для аріорного моделювання атак та їх наслідків використовується численна кількість методів і, не зважаючи на те, що такі моделі охоплюють і атаки і їх наслідки, жодна з них не забезпечує оцінку впливу атаки при її незалежному використанні. Щоб позбутися цього недоліку, необхідно побудувати комплексну модель, яка б дозволяла точно оцінити вплив багатоступінчастості для виявлення шляхів відбиття атаки з високою ефективністю. Задначена комплексна модель може бути основою для парадигми Ситуаційної Кіберобізнаності, яка використовується аналітиками при прогнозуванні наслідків атак. При цьому, основними елементами моделі мають стати: карта кроків атаки, що використовує наші уразливості; системи та сервіси, які можуть бути уражені; дотичні системи, які також можуть бути уражені.

Аналіз джерел

Сценарії атак та стратегії їх здійснення на основі кореляційних залежностей розглядаються у [1]. Інші технології поєднують джерела атаки та їх параметри на основі статистичних залежностей [2]. У [3] застосовується підхід на основі кореляції уразливостей зі сповіщеннями, оскільки він може ефективно відфільтровувати хибні попередження на основі процедур вкладених циклів.

Для вирішення головної проблеми найбільш широко використовуються напів-Марківські моделі та їх варіанти. Наприклад у [4] запропоновано модель, що є дворівневим розширенням прихованої напів-Марківської моделі. Крім того можуть використовуватись динамічні Баєсівські мережі [5] та ймовірнісні розширення мереж Петрі [6] для контролю та розпізнавання мультиагентної активності. Також придатними можуть бути результати досліджень щодо управління потоками данних, які можуть використовуватись для кореляції потоків з вузлами графа атаки. Великі мережі підприємств можуть залежати від багатьох хостів та компонентів. У [7] запропоновано дослідження параметрів та обмежень цього класу залежностей з використанням інформації про пакети та час мережевого трафіку.

Таким чином, можливості точного моделювання різного класу залежностей в комплексних мережах та оцінка впливу багатокрокових атак є чутливими до числа застосувань, включаючи поведінкові підходи. У той же час створення комплексної моделі багатоступеневої атаки на основі поєднання уразливостей та способів її здійснення залишається широким полем для наукових досліджень.

Метою даної статті є розробка графових моделей кібербезпеки, які б забезпечували можливість оцінювання можливих сценаріїв атаки та проектування засобів захисту відносно конкретної топології мережі.

Виклад основного матеріалу

Архітектура мережі. Для побудови комплексної моделі за основу розгляду взято мережу (рис. 1) [8], яка включає 3 підмережі, розділені міжмережевими екранами. Дві з трьох підмереж надають інтернет-сервіси, наприклад “онлайн-магазин” та “трекінг замовлень”, перший з яких розгорнуто на веб-сервері H_A, сполученому з сервером локальної бази даних H_B, в той час, як другий сервіс запроваджено на сервері мобільних застосувань H_C, який сполучено з сервером локальної бази даних H_D. З рештою, логіка роботи системи впроваджується у підмережі, яка включає 2 сервери ядра застосувань H_E та H_F та центральний сервер бази даних H_G, який підтримує інформацію щодо продукції, споживачів та замовлень. Два локальних сервери баз даних діють як локальний кеш для відповідних підмереж. Ідеальним сценарієм атаки у такому випадку буде ураження сервера центральної бази даних та знищення важливої фінансової інформації.

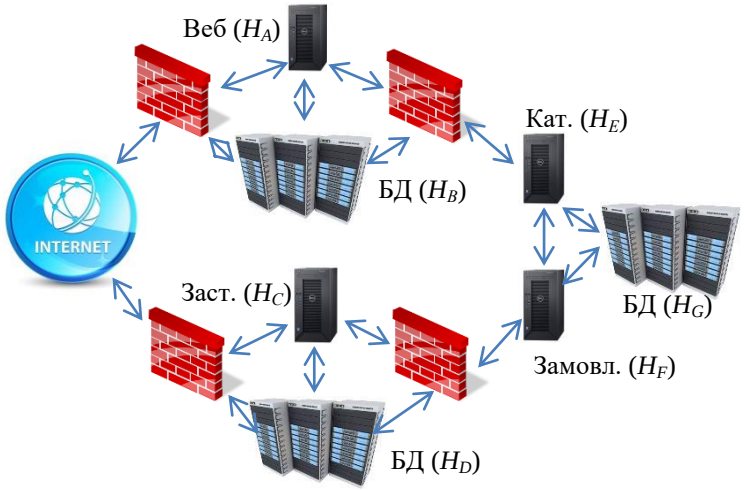

Рис. 1. Мережева архітектура онлайн магазину

Для аналізу змоделюємо спочатку топологію мережі, відомі уразливості, включаючи уразливості нульового дня та їх взаємозалежності. Уразливості завжди взаємозалежні і тому аналіз їх окремо не є надто ефективним. Топологічний підхід до аналізу уразливостей дозволяє побудувати граф атаки, подаючи усі можливі шляхи атаки мережевим графом. Вузли у такому графі представляють, у залежності від рівня абстрагування, уразливості підмережі чи окремого хоста, чи, можливо, окремого програмного застосування. Дуги представляють випадкові відносини між уразливостями. Далі необхідно зробити аналіз залежностей між сервісами та хостами для побудови графа залежностей. Аналіз залежностей є критичним для оцінювання поточного впливу у залежності від атаки, що триває. Дві зазначені моделі можуть бути інтегровані у загальний багатошаровий граф оцінки впливу.

Відображення шляхів атаки за допомогою топологічного аналізу забезпечує розуміння того, як індивідуальні та комбіновані уразливості впливають на загальну безпеку мережі. Наприклад, так можна визначати:

як окремі заходи щодо зниження ризику впливають на загальну безпеку;

як нова уразливість впливає на загальну безпеку мережі;

які зміни у безпеці окремого хоста можуть мати вплив на загальний ризик для підприємства.

Граф атаки. Першим кроком для ефективної оцінки впливу багатокрокової атаки є встановлення можливих наслідків атаки. Граф атаки відображає апріорне бачення

уразливостей у їх залежності від топології мережі та призначається для врахування усіх можливих слабких місць, які може використати атакуючий.

Математичний запис графа атаки може бути наведено наступним чином.

Граф атаки – це спрямований граф $G=(V\cup C, R_r\cup R_i)$, у якому $V\cup C$ – множина вершин, $R_r\cup R_i$ – множина дуг. При цьому V – множина експлойтів уразливостей, C – множина умов безпеки, $R_r\subseteq C\times V$ – множина запитів та $R_i\subseteq V\times C$ – множина відповідей (реакції системи захисту).

Експлойт визначається як предикат $v(H_S,H_D)$, який вказує, що використовуєтсья уразливість v на цільовому хості H_D з хоста атаки H_S. Стан безпеки описується предикатом $c(H_S,H_D)$, який показує умови безпеки c, що включають хост-джерело H_S та цільовий хост H_D.

У такому випадку початкові умови можуть бути визначені як множина $C_i=\{c\in C|\exists v\in V,(v,c)\in R_i\}$, у той час як проміжні (поточні) умови можуть визначатися як $C\setminus C_i$. Проміжні умови – це результати застосування експлойтів і тому вони не можуть бути усунуті без усування причини їх виникнення, у той час як початкові умови можуть бути змінені.

Граф залежності уразливостей. Оскільки метою роботи є оцінювання впливу, де експлойти відіграють центральну роль, то є доцільним ввести граф залежності уразливостей $G=(V,R)$, де $R=\{(v_i,v_j)\in V\times V|\exists c\in C,(v_i,c)\in R_i\wedge(c,v_j)\in R_r\}$ є множиною вузлів. Оскільки передумова до атаки повинна бути пов'язана з певною уразливістю і кожна така передумова передбачає, що хоча б один з експлойтів буде успішним, то такі відносини описуються кон'юнкцією диз'юнкцій.

Ймовірнісний граф атаки. Для опису моделей атак широко використовуються графи атаки. Разом з тим, для багатоступеневих атак, коли одночасно або за певним розкладом виконується велика кількість операцій, такий граф потребує удосконалення, зокрема шляхом введення часового розподілу. Передбачається, що кожен крок атаки виконується в межах певного фрейму після того, як попередній крок був реалізований. При цьому кожен з них асоціюється з певним значенням ймовірності.

Часовий розподіл. Оцінювання середнього часу здійснення атаки можна здійснити на основі часового розподілу $\omega=(I,\rho)$, де I – множина часових інтервалів, таких, що $\forall[x,y]\in I, x\le y$; $\forall[x,y],[x',y']\in I$ таке, що $[x,y]\ne[x',y']$, інтервали $[x,y]$ та $[x',y']$ є відокремленими; $\rho(x,y)\in[0,1]$ – функція, яка асоціюється з часовим інтервалом $[x,y]\in I$.

Часовий розподіл (I,ρ) визначає множину I відокремлених часових інтервалів, коли окремий експлойт може бути застосований та розподіл ймовірностей $\rho(x,y)\in[0,1]$ на I, що показує ймовірність того, що експлойт буде застосований на проміжку $[x,y]$.

Базуючись на визначенні графа атаки $G=(V\cup C, R_r\cup R_i)$, ймовірністний граф атаки може бути визначено як позначений спрямований ациклічний граф $A=(V,E,\delta,\gamma)$, де V – скінченна множина експлойтів у графі атаки; $E=R_i\circ R_r$; $V^s=\{v\in V|\exists v'\in V,(v',v)\in E\}\ne\varnothing$ і $V^e=\{v\in V|\exists v'\in V,(v,v')\in E\}\ne\varnothing$ тобто існує хоча б один початковий і один кінцевий вузол; $\delta: E\to\Omega$ функція часового розподілу кожного вузла графа, така що $(\forall v\in V)\sum\limits_{\{v\in V|(v,v')\in E\}}S(\delta(v,v'))=1$; γ – функція, яка пов'язана з кожним експлойтом $v_j\in V\setminus V^s$ та визначається як $\gamma(v_j)=\bigwedge\limits_{c_k\in C|(c_k,v_j)\in R_r}\left(\bigvee\limits_{v_i\in V|(v_i,c_k)\in R_i}v_{i,j}\right)$, де $v_{i,j}$ визначає, що v_i передує v_j та

виконується в одному з часових інтервалів $\delta(v_i, v_j)$, який слідує за виконанням v_i. Для формалізації стану атаки доцільно застосувати дерево $T = (V_T, E_T)$ на графі $A = (V, E, \delta, \gamma)$ таке, що $|V_T \cap V^e| = 1$ – є хоча б один кінцевий вузол на графі *T*; $(\forall v \in V_T \setminus V^s) \exists V' \subset V_T$ таких, що V' задовольняє $\gamma(v) \wedge \nexists V'' \subset V'$.

Граф залежності. Сучасні інформаційні системи складаються з великого числа взаємозалежного програмного забезпечення та хостів. Відтак, збій у роботі одного з них може спричинити ланцюгову реакцію, яка призведе до виходу з ладу усієї мережі. Тому, для подальшої формалізації задач необхідно ввести моделі залежностей окремих компонентів мережі, зокрема:

надмірність – мережевий компонент залежить від множини надмірних компонентів;

деградація – мережевий компонент залежить від множини інших компонентів так, що при виході з ладу одного, система може працювати поступово втрачаючи можливості;

повна залежність – мережевий компонент залежить від множини інших так, що коли один виходить з ладу, то всі залежні компоненти перестають функціонувати.

Для подальшого розвитку моделі необхідно прийняти, що кожен компонент (обладнання чи програмний засіб) може характеризуватися показником доступності за шкалою від 0 до 1, де 1 означає, що компонент повністю доступний, а 0 – недоступний. Також доцільно ввести функцію, яка буде показувати на скільки доступним є компонент у залежності від інших взаємодіючих компонентів. Така функція, це *функція залежності*, при чому $f(0,...,0) = 0$ та $f(1,...,1) = 1$.

Тоді, відповідно до наведених сценаріїв можна буде записати сукупність рівнянь, які відповідатимуть різним станам мережевих компонентів у залежності від стану інших.

У випадку надмірності $f_r(l_1,...,l_n) = \begin{cases} 1, \textit{якщо } \exists i \in [1,n], l_i = 1; \\ 0, \textit{у протилежному випадку}. \end{cases}$

У випадку деградації $f_d(l_1,...,l_n) = \frac{1}{n}\sum_{i=1}^{n} l_i$.

При повній залежності $f_s(l_1,...,l_n) = \begin{cases} 1, \textit{якщо } \forall i \in [1,n], l_i = 1; \\ 0, \textit{у протилежному випадку}. \end{cases}$

Узагальнений граф залежності. Узагальнений граф залежності – це позначений спрямований ациклічний граф $D = (H, Q, \phi)$, де *H* – множина вузлів, повязаних з мережевими компонентами; $Q = \{(h_1, h_2) \in H \times H | h_1 \textit{ залежить від } h_2\}$ – множина дуг; ϕ – відображення, яке пов'язує вузли *H* з фунцкціями *f*. Для кожного вузла $h \in H$, *h* визначає множину компонентів, які залежать від *h*.

Функція мережевого статусу. Базуючись на визначенні узагальненого графа залежності функція мережевого статусу для $D = (H, Q, \phi)$ може бути визначена як функція $s: H \times T \to [0,1]$ така, що $\forall h \in H$ та $\forall t \in T$, $s(h,t) \le f(s(h_{i_1}, t),..., s(h_{i_m}, t))$, де $f = \phi(h)$, $h = \{h_{i_1},..., h_{i_m}\}$ – множина компонентів від яких залежить *h*. Доступність компонента *h* пов'язана з відповідною функцією залежності, яка, теоретично, є максимальною. На практиці доступність може бути нижчою у випадку безпосередньої компрометації компонента *h* у результаті атаки.

Комбінована графова модель може бути побудована шляхом інтеграції графів атаки та залежності у загальну багатошарову модель з метою оцінювання впливу триваючих атак у тому числі і перспективних.

Повертаючись до прикладу, наведеного на початку статті, мережа, яка надає онлайн сервіси (торгівля та трекінг замовлень), складається з трьох підмереж розділених мережевими екранами. Перші дві підмережі включають хости з доступом до мережі

Інтернет, третя підмережа застосовує ключові бізнес-логіки та використовує центральний сервер бази даних. Атакуючий, який хоче одержати важливі дані з головного сервера бази даних повинен буде проникнути через декілька фаєрволів та одержати доступ до хостів, перш ніж досягнути мети.

На рис. 2 наведено ймовірнісний граф атаки для мережі рис. 1. Тут видно, що як тільки уразливість v_C на хості h_C використана, ми можемо зпрогнозувати, що атакуючий використає уразливість v_D на хості h_D, або v_F на хості h_F. Ймовірнісна природа цієї моделі дає можливість визначити, яким чином буде розвиватися атака. У той же час граф атаки сам по собі не дає відповіді на важливі питання: яка атака все ж таки матиме більший вплив на сервіси, які забезпечуються мережею? Як можна зменшити ризик? Модель, яка розробляється, повинна дати відповіді на ці питання.

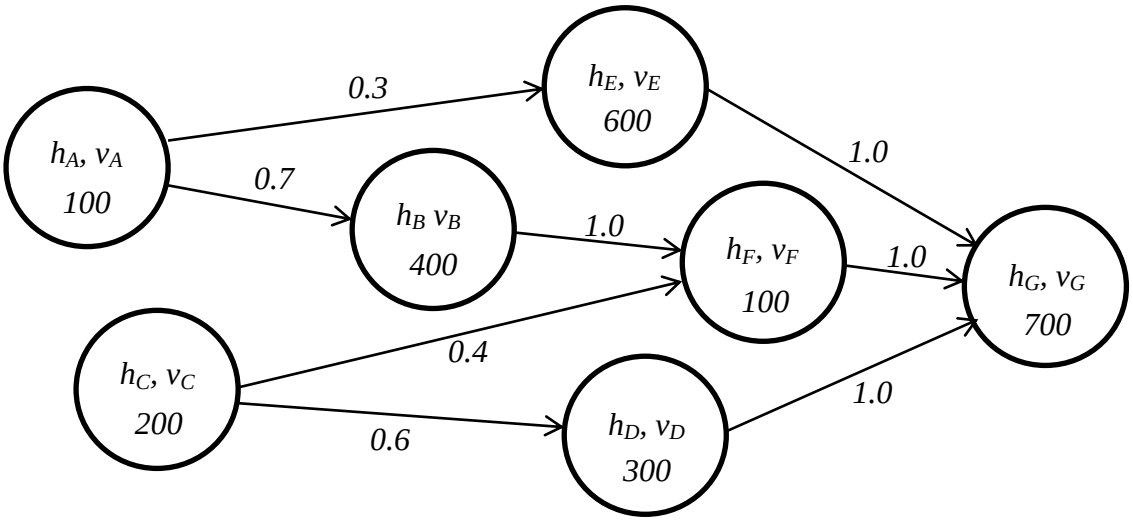

Рис. 2. Комбінований ймовірнісний граф атаки

На перший погляд, атака буде розвиватися за більш ймовірним сценарієм $h_A \to h_B \to h_F \to h_G$. У той же час, урахування вартості ресурсу, який уражається, може дати іншу картину, що і визначатиме необхідні заходи захисту. Для цього необхідно оцінити всі можливі варіанти атаки:

$h_A \to h_B \to h_F \to h_G$, 100 + 280 + 100 + 700 = 1180;

$h_A \to h_E \to h_G$, 100+180+700 = 980;

$h_C \to h_F \to h_G$, 200 + 400 + 700 = 1300;

$h_C \to h_D \to h_G$, 200 + 180 + 700 = 1080.

Таким чином, як бачимо, шлях розвитку атаки $h_C \to h_F \to h_G$ буде мати важчі наслідки для системи, ніж більш ймовірний $h_A \to h_B \to h_F \to h_G$. Тому, врахування у графовій моделі як ймовірностей так і вартості ресурсів дає можливість більш адекватно оцінювати можливі шляхи розвитку атаки та їх наслідки.

Висновки

Застосування графових моделей для прогнозування та оцінювання наслідків кібератак дає можливість наочно оцінити ступінь уразливості системи, можливі шляхи розвитку складних багатоступінчастих атак та встановити можливі варіанти захисту. Такий підхід є перспективним для широкого кола організацій та установ з огляду на особливості їх функціонування. Основою для визначення параметрів графів можуть бути як реальні показники діяльності організації, так і результати експертних оцінок.

Напрямком подальших досліджень можуть бути різноманітні аспекти удосконалення графових моделей кібербезпеки, які б інкапсулювали можливості експертного та технічного рівнів захисту.

Перелік посилань

1. Сердюк В.А. Анализ современных тенденций построения моделей информационных атак / В.А. Сердюк // Информационные технологии. – 2004. – № 5. – С. 94–101.

2. Абрамов Е.С., Андреев А.В., Мордвин Д. В. Применение графов атак для моделирования вредоносных сетевых воздействий // Известия Южного федерального университета. Технические науки. – 2012, № 5. – С. 165–173.

3. Колегов Д.Н. Проблемы синтеза и анализа графов атак. https://www.securitylab.ru/contest/299868.php.%202007?R=1.

4. Duong T.V., Bui H.H., Phung D.Q., et al. Activity recognition and abnormality detection with the switching hidden semimarkov model // In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2005), volume 1. San Diego, CA, USA, pp. 838–845.

5. Коломеец М.В. Чечулин А.А. Дойникова Е.В. Котенко И.В. Методика визуализации метрик кибербезопасности // Изв. вузов. Приборостроение. – 2018. – Т. 61, № 10. – С. 873–879.

6. Albanese M., Chellappa R., Moscato V., et al. A constrained probabilistic petri net framework for human activity detection in video.IEEE Transactions on ultimedia. – 2008; 10(8): 1429–1443.

7. Kotenko I.V., Chechulin A.A. The use of attack to evaluate the security of computer networks and analysis of security events [Primenenie grafov atak dlya otsenki zashchishchennosti komp'yuternykh setey i analiza sobytiy bezopasnosti], Sistemy vysokoy dostupnosti [High Availability Systems]. – 2013, Vol. 9, no. 3. – pp. 103–110.

8. Massimiliano Albanese and Sushil Jajodia. A Graphical Model to Assess the Impact of Multi-Step Attacks. Journal of Defense Modeling and Simulation: Applications, Methodology, Technologyб 2018, Vol. 15(1). – P. 79–93.

Надійшла: 09.09.2019 Рецензент: д.т.н., професор Вишнівський В.В.