РОЗДІЛ ІІ. ІНФОРМАЦІЙНО-КОМП'ЮТЕРНІ ТЕХНОЛОГІЇ

УДК 004.056.5:004.7(083.94)
DOI: 10.25140/2411-5363-2021-1(23)-53-61

Сергій Толюпа, Іван Пархоменко, Людмила Терейковська, Володимир Квасніков

ПОБУДОВА СИСТЕМ ВИЯВЛЕННЯ КІБЕРАТАК ЗА ДОПОМОГОЮ ПРИХОВАНОЇ МАРКІВСЬКОЇ МОДЕЛІ

Одним із найбільш перспективних напрямків підвищення якості аналізу даних є використання їх у системах виявлення мережевих кібератак методом виявлення аномалій. У цьому методі робота аналізаторів даних базується на припущенні, що ознакою кібератаки служить деяке відхилення контрольованих параметрів комп'ютерної системи від параметрів, що характеризують нормальне функціонування мережі. У результаті проведених досліджень обґрунтована можливість формування шаблонів нормальної поведінки мережевих об'єктів комп'ютерних систем на основі однорідного ланцюга Маркова з послідовними переходами.

Ключові слова: *мережева кібератака; комп'ютерна система; інформаційна безпека; система виявлення мережних кібератак; ланцюг Маркова; марковський процес; марковська модель; характеристика безпеки.*
Рис.: 4. Бібл.: 11.

Актуальність теми дослідження. Одним з найбільш перспективних напрямів підвищення якості аналізу даних є використання їх у системах виявлення мережевих кібератак (СВА) методом виявлення аномалій. У цьому методі робота аналізаторів даних базується на припущенні, що ознакою кібератаки служить деяке відхилення контрольованих параметрів комп'ютерної системи (КС) від параметрів, що характеризують нормальне функціонування мережі. При цьому значення контрольованих параметрів за звичайних умов експлуатації отримали назву шаблонів нормальної поведінки. Саме тому дуже важливим завданням є вибір або формування такого шаблону, який би адекватно відтворював функціональний портрет мережевого об'єкта КС і дозволив із заданою точністю визначити аномальну поведінку цього об'єкта.

Постановка проблеми. Різні типи кібератак, які можуть мати масований характер дії на об'єкти інформаційних систем, спонукають до виникнення певних технічних рішень, що мають спеціалізоване спрямування. Для знаходження мережевої активності, яку можна ідентифікувати як мережеве вторгнення, застосовують певні методи [1-5], спеціальні моделі та технічні рішення для систем, які виявляють та запобігають вторгненням [6; 7; 8; 11]. І цей комплекс підходів і рішень може використовуватися і проти інших типів кіберзагроз, які з'являються і мають відмінні чи модифіковані характеристики. У разі появи модифікованих чи відмінних загроз, які генеруються певними атакуючими діями і мають невизначені характеристики, можуть бути не виявленими при використанні зазначених засобів, які можуть потребувати суттєвих часових затрат для адаптації під відповідні типи кіберзагроз. Для ефективного функціонування систем, які застосовуються для виявлення вторгнень (СВВ), необхідне неперервне їх удосконалення та модифікація під нові типи кіберзагроз.

Аналіз останніх досліджень і публікацій. На сьогодні вирішення питань забезпечення безпеки в інформаційних системах (ІС) та управління станом їх захищеності описується в роботах вітчизняних та закордонних дослідників, а саме: В. Бурячка, С. Гнатюка, О. Корченко, О. Кузнецова, І. Субача, С. Євсеєва, В. Дудикевича, С. Казмирчук, І. Тейраковського, T. Ptaceka, G. Elmasry, P. Albers, O. Camp та інших.

Безперечно, актуальним напрямом у сфері інформаційної безпеки, який має інтенсивний розвиток, є напрям з виявлення кібератак та захисту від втручань у роботу інформаційних систем з боку неавторизованих сторін. Крім того, слід зазначити, що атаки на інформаційні системи відбуваються дедалі частіше, методи їх реалізації більш досконалі, а масштаби – більш глобальні.

Виділення недосліджених частин загальної проблеми. Сучасні системи виявлення вторгнень і атак ще далекі від ергономічних і ефективних з погляду безпеки рішень. Підвищення ефективності ж слід вести не тільки в області виявлення зловмисних дій на інфраструктуру захищених об'єктів інформатизації, але і з погляду повсякденної експлуатації цих засобів, а також економії обчислювальних та інформаційних ресурсів власника цієї системи захисту. Використання одного з ефективних методів виявлення вторгнень та атак основується на сигнатурному підході. Сигнатурні методи дозволяють описати атаку набором правил або за допомогою формальної моделі, у ролі якої може застосовуватися символьний рядок, семантичне вираження спеціальною мовою тощо. Суть цього методу полягає у використанні спеціалізованої бази даних шаблонів (сигнатур) атак для пошуку дій, що підпадають під визначення «атака».

Метою статті є підвищення якості аналізу даних у системах виявлення мережевих кібератак методом виявлення аномалій за допомогою прихованих марківських моделей.

Виклад основного матеріалу. Прихована марковська модель має модельовану систему, яка вважається марковським процесом із прихованими станами, тобто зі станами, які важко піддаються контролю. Прихована марковська модель є статичною марковською моделлю і вона може представлятися у вигляді динамічної баєсової мережі. При використанні ланцюгів Маркова (простішої марковської моделі) для спостерігача буде видимий стан. І в такому випадку єдиними параметрами є ймовірності переходу стану. Але якщо взяти приховану марковську модель, то в цьому випадку неможливо безпосередньо спостерігати стан. Хоча буде видно вихід, що залежить від стану. У цьому випадку вихід може бути у вигляді даних «токена». І залежно від вихідних токенів кожний стан буде мати розподіл ймовірностей. Отже, можна отримати певну інформацію про послідовність станів, знаючи послідовність токенів, яка генерується прихованою марковською моделлю. Саме до послідовності станів, через яку дана модель проходить, а не параметрів цієї моделі відноситься поняття «приховане». Це так навіть і в тому випадку, якщо ймовірності переходу (ці параметри) відомі.

Використовуючи цей підхід, на відміну від існуючих систем виявлення вторгнень, при роздільній обробці різних етапів атаки отримуємо можливість розпізнати загрозу не в стадії реалізації, а в процесі її формування. В такому випадку основою, що реалізує даний підхід і забезпечує розпізнання загроз може бути сигнатурний пошук, чи механізми виявлення аномалій. Крім того буде доречним і використання експертних систем та методів. Вже відомих локальних і мережних примітивів оцінки потоку подій, які відбуваються в інформаційному середовищі.

Під час побудови СВА доцільно врахувати загальну структуру побудови системи управління подіями та вразливостями безпеки, до яких входять як внутрішні, так і зовнішні загрози. У свою чергу, зважаючи на проведений аналіз, для побудови СВА від внутрішніх атак доцільно використовувати метод опорних векторів, а для побудови СВА від зовнішніх атак – метод динамічного програмування. Сутність цих кроків полягає в ідентифікації станів захищеності з урахуванням множини різнорідних параметрів трафіка, що передається в ІС.

Структурно-логічну схему системи управління подіями, інцидентами, атаками інформаційної безпеки в ІС із застосуванням інтелектуальних технологій представлено на рис. 1.

При виявленні кібератак існують два загальні підходи:

- розпізнавання сигнатур (шаблонів);
- виявлення аномалій, які доповнюють один одного при виявленні кібератаки.

Розглядаючи сигнатурний метод для виявлення вторгнень та атак на інформаційну систему, можна зазначити що його ефективність залежить від якості спеціалізованої бази сигнатур (певних шаблонів) атак, які в свою чергу забезпечують пошук дій, що ідентифікуються як атаки. При сигнатурному підході атака описується певним набором правил чи використовується формальна модель у вигляді семантичного вираження на спеціальній мові.

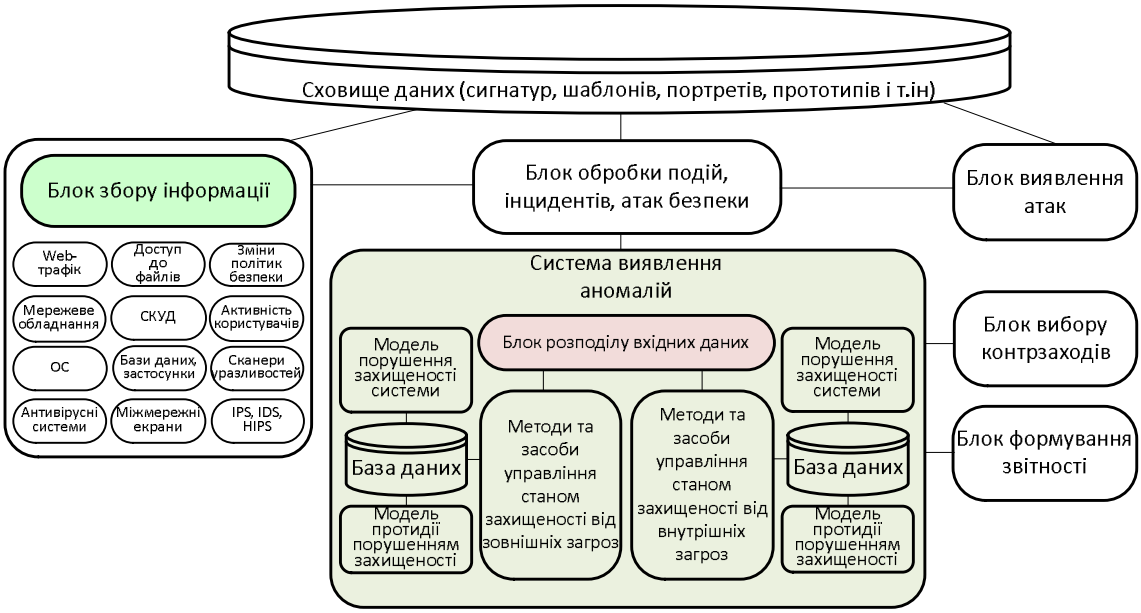

Рис. 1. Структурна схема системи управління інформаційною безпекою

Здебільшого цей метод ефективний у тому випадку, коли відомий фрагмент коду вірусу, тобто сигнатура вірусної чи хакерської атаки і цей код внесений у базу даних системи виявлення вторгнень. Отже, коли проти інформаційної системи вперше застосували певний тип хакерської атаки, чи ця система заражена якимсь новим видом вірусу, то, звісно, в базі немає відповідних шаблонів (сигнатур). І в цьому випадку система виявлення вторгнень на базі сигнатурного підходу не повідомляє про загрозу, бо не вважає атакуючі впливи небезпечними.

Отже, ефективність використання системи виявлення вторгнень на базі сигнатурного підходу залежить від таких факторів: оперативності та вчасності додання нових сигнатур, наповненість і повнота шаблонів атак та використання спеціальних інтелектуальних алгоритмів ідентифікації зловмисних дій на підставі виділення певних базових кроків, використовуючи порівняння із сигнатурами.

Що стосується модулів обробки даних, то в цьому випадку в системі обробки інформації кожна сигнатура, шаблон атаки є базовим компонентом для визначення фаз атаки та етапів реалізації. Отже, сигнатура ідентифікує певні загальні дії та її поняття відповідає деякому вирішальному правилу, а в конкретній атаці виділяються окремі етапи її проведення. Звичайно для простих атак інструменти аналізу дають більше можливостей для їх виявлення.

Етапи реалізації атаки можуть бути представлені графом переходів. Що стосується фаз атаки, то вони можуть бути такими: перевірка портів; визначення програмно-апаратних засобів; використання вразливостей у програмному забезпеченні (за допомогою експлойтів); проведення DoS чи DDoS атак для дезорганізації персоналу; використання несанкціонованого віддаленого доступу до комп'ютера (за допомогою бекдорів); виявлення троянів, які вже встановленні в системі; виявлення proxy-серверів; знищення цифрових слідів у системі і т. ін. Деталізація фаз атак може бути різною.

Існують такі проблеми використання прихованих марковських моделей для виявлення аномалій:

- проблема оцінки ймовірності того, що послідовність, яка спостерігається, була згенерована моделлю;
- проблема побудови із даних аудиту моделі або набору моделей, які б правильно описували поведінку, яка спостерігається;
- проблема обчислення найбільш ймовірного набору прихованих станів, що призвели до отримання спостереженням у межах прихованої марковської моделі.

У разі виявлення мережевих аномалій контроль трафіку формує дані, які можуть бути використані для навчання прихованої марковської моделі. Модель, побудована на основі аналізу поведінки цільової системи, є відповідним поданням для профілю трафіку.

У роботі [1] прихована марковська модель для нормальної поведінки формується з журналів і даних нормальної поведінки мережної системи. Поведінка системи, яка спостерігається, аналізується для оцінки ймовірності того, що побудована модель підтримує поведінку, яка спостерігається. Низька ймовірностей підтримки вказує на аномальну поведінку.

У роботі [2] описується застосування прихованої марківської моделі для виявлення аномалій з використанням профілю послідовностей системних викликів і shell-команд. При навчанні модель обчислює ймовірність вибірки послідовності, яка спостерігається. Граничні значення ймовірності формуються на основі мінімальної ймовірності серед усіх послідовностей у даних, які навчаються та використовуються для розмежування між нормальною та аномальною поведінкою. Головним обмеженням цього підходу є слабке узагальнення, що може призвести до блокування користувачів, неоднозначно визначених у цій системі.

У багатьох випадках працездатність ІС можливо оцінити за величиною одного або декількох параметрів, контрольованих на експлуатації. Здебільшого для кожного із зазначених параметрів можливо визначити специфічну область значень (область працездатності), вихід з якої відповідає рівню параметричної відмови ІС. При цьому працездатний стан ІС відповідає перебування всіх контрольованих параметрів в своїх областях працездатності [7]. Будемо вважати, що відомий повний перелік цих параметрів, крім цього, всі вони контролюються. Відзначимо, що в цей перелік параметрів доцільно поділити на дві групи. До першої групи слід віднести параметри, які безпосередньо характеризують працездатність ІС, наприклад, обсяг оперативної пам'яті, що використовується, або завантаження центрального процесора. До другої групи слід включити параметри, що дозволяють побічно оцінити працездатність КС, наприклад, обсяг черги запитів до комп'ютера сервера.

Як показує практичний досвід, більшість параметрів, що характеризують працездатність КС, специфічно змінюються при здійсненні мережевої кібератаки. Причому атака на відмову в обслуговуванні може змінювати параметри обох груп, а атака з метою отримання несанкціонованого доступу переважно впливає на параметри другої групи. Таким чином, специфічна зміна параметрів другої групи сигналізує як про здійснення кібератак на відмову в обслуговуванні, так і про здійснення кібератак з метою отримання несанкціонованого доступу. Для визначення зазначеної специфіки доцільно провести моделювання зміни параметрів другої групи на певному інтервалі часу $t \in [0, t_{\max}]$ як за нормальних умов експлуатації, так і при здійсненні атаки на КС. Основою моделювання може стати марковська модель апроксимації динаміки контрольованих параметрів. У разі застосування нестаціонарної марковської моделі розрахунок імовірнісних характеристик динаміки параметрів можна виконати шляхом вирішення системи рівнянь Колмогорова-Чепмена:

$$\begin{cases} P_1(t) = P_1(t-1) - P_1(t-1)\eta_{1,i} + P_1(t-1)\eta_{1,N} \\ P_i(t) = P_i(t-1) - P_i(t-1)\eta_{i,j} + P_i(t-1)\eta_{i,N} \\ P_N(t) = P_N(t-1) + P_i(t-1)\eta_{i,N} + P_{N-1}(t-1)\eta_{N-1,N} \end{cases}, \quad (1)$$

де $P_i(t)$ – ймовірність перебування параметра в i-му кванті в момент часу $t \in [0, t_{\max}]$, $\eta_{i,j}$ – інтенсивність переходу з кванта i в квант j.

У цьому випадку під інтенсивністю переходу розуміється ймовірність переходу з одного кванта в інший квант за одиницю часу.

На наш погляд, побудова загальної моделі виду (1) недоцільно як з погляду труднощів, що виникають при отриманні та підготовці вихідних даних для моделювання (інтенсивностей переходу), так і з погляду, складності інтерпретації результатів моделювання. Розглянемо можливі шляхи спрощення моделі.

Відповідно до реальних можливостей отримання результатів контролю параметрів другої групи відповідно до [7; 10] як модель шаблону доцільно застосувати однорідну марковську модель із можливістю тільки послідовних переходів по станам. Розмічений граф такої моделі показаний на рис. 2.

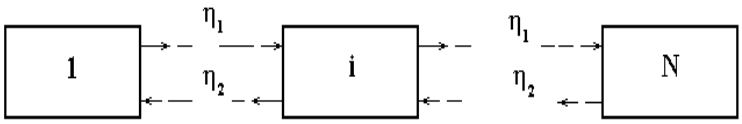

Рис. 2. Розмічений граф однорідної марковської моделі з послідовними переходами

На рис. 3 кількість квантів марковского процесу позначено як N, інтенсивність переходу з кванту i в квант i+1 марківського процесу позначено як η_1, а інтенсивність переходу із кванту i+1 в квант i позначено як η_2. Як видно із рис. 1, можливі тільки послідовні переходи по квантах марковського ланцюга, при цьому справедлива умова:

$$\begin{cases} \eta_{1,2} = \eta_{2,3} = ... = \eta_{i,,i+1} = ... = \eta_{N-1,N} = \eta_1 \\ \eta_{2,1} = \eta_{3,2} = ... = \eta_{i,,i-11} = ... = \eta_{N,N-1} = \eta_2 \end{cases} \quad (2)$$

Зазначені передумови дозволяють істотно спростити математичний апарат розрахунку перебування параметра в кванти марковського процесу:

$$\begin{cases} P_1(t) = P_1(t-1) - P_1(t-1)\eta_1 + P_2(t-1)\eta_2 \\ P_i(t) = P_i(t-1) - P_i(t-1)\eta_1 + P_{i-1}(t-1)\eta_2 \\ P_N(t) = P_N(t-1) - P_N(t-1)\eta_2 + P_{N-1}(t-1)\eta_1 \end{cases} \quad (3)$$

У теорії експлуатації технічно складних систем подібні моделі застосовуються при моделюванні зміни технічного стану відновлюваних агрегатів, для яких характерна двостороння область працездатності. У роботі [7] показана можливість спрощення цієї моделі за рахунок переходу до агрегатів, що не відновлюються з обмеженою зверху односторонньої областю працездатності. При цьому як марковська модель використовується однорідний марковській ланцюг, що поглинається з постійною інтенсивністю переходів, що дозволяє спростити систему рівнянь (3) таким чином:

$$\begin{cases} P_1(t) = P_1(t-1) - P_1(t-1)\eta \\ P_i(t) = P_i(t-1) - P_i(t-1)\eta + P_{i-1}(t-1)\eta \\ P_N(t) = P_N(t-1) + P_{N-1}(t-1)\eta \end{cases} \quad (4)$$

При вирішенні (4) застосовується умова нормування:

$$\sum_{i=1}^{N} P_i(t) = 1 \quad (5)$$

Умови нормування вказують на те, що в будь-який момент часу значення параметру може знаходитися тільки в одному із заздалегідь відомих певних квантів.

Початкові умови моделювання припускають знаходження параметра, що досліджується, в нульовий момент часу в початковому стані.

$$P_1(0) = 1 \quad (6)$$

Результати [7; 10] дозволили визначити, що оптимальна кількість квантів марковського процесу *N* дорівнює 20. Така кількість квантів забезпечує достатню точність моделювання, при прийнятному використанні обчислювальних ресурсів. Величини квантів рівні між собою. Застосовується безрозмірна обмежена зверху одностороння область працездатності. У цьому випадку параметр, що визначає працездатність КС, може змінюватися в межах від 0 до 1. Також прийнято припущення, що часовий інтервал експлуатації, а отже, і часовий інтервал моделювання [0, 1].

Розроблена математична модель дозволяє розрахувати ймовірність перебування параметра в певних квантах залежно від величини інтенсивностей переходу й поточного напрацювання (часу експлуатації) досліджуваної КС. Результати такого розрахунку дозволяють визначити математичне очікування параметра й показники надійності КС. На рис. 3 показані графік залежності математичного очікування досліджуваного параметра τ від часу $M^{\tau}=f(t)$, а також графіки залежності ймовірності перебування досліджуваного параметра τ в квантах марковського процесу від часу $P^{\tau}=f(t)$ при інтенсивності переходу марковського процесу $\eta=0{,}02$.

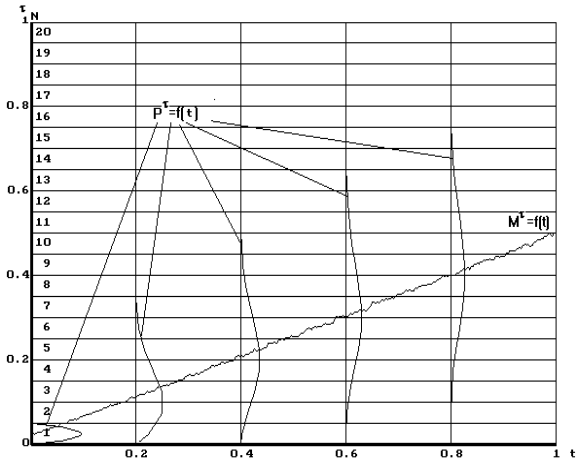

Рис. 3. Розмічений граф однорідної марковської моделі з послідовними переходами

Проведено моделювання динаміки досліджуваного параметра для інтенсивностей переходу $\eta \in \{0.02, 0.04, 0.06, 0.08, 0.1\}$.

Відповідні графіки залежностей математичного очікування досліджуваного параметра від часу експлуатації показано на рис. 4.

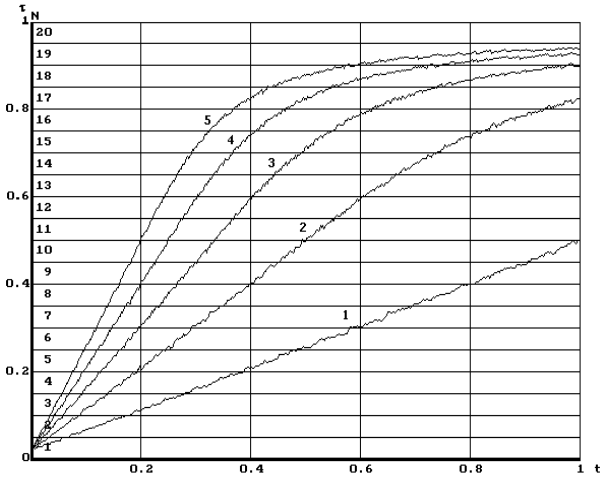

Рис. 4. Графіки математичного очікування параметра від часу експлуатації

Висновки. Побудова подібних залежностей для різних інтенсивностей переходу може стати основою для формування профілів нормальної поведінки КС. У результаті проведених досліджень обґрунтована можливість формування шаблонів нормальної поведінки мережевих об'єктів комп'ютерних систем на основі однорідного ланцюга Маркова з послідовними переходами. Розроблено структуру моделі, сформовано відповідне математичне забезпечення.

Також слід зазначити, що в багатьох випадках оцінити ймовірність кібератаки на КС за значенням одного, нехай навіть узагальнюючого параметра, досить важко. Зазвичай здійснення кібератаки пов'язано із синхронною зміною декількох параметрів. Таким чином, важливим напрямком подальших досліджень є розробка модельних залежностей спільної зміни параметрів, що визначають безпеку КС. Крім цього, розробка шаблонів нормальної поведінки КС недоцільна без обґрунтування номенклатури та методу оцінки контрольованих параметрів. Отже, ще одним важливим напрямом подальших досліджень має стати розробка системи вибору та оцінки параметрів безпеки КС.

Список використаних джерел

1. Tereikovskiy I, Toliupa S., Parkhomenko I., Tereikovska L. «Markov Model of Normal Conduct Template of Computer Systems Network Objects». *14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering TCSET-2018,* pp. 498–501.

2. Aitchanov B., Korchenko A., Tereykovskiy I., Bapiyev I. Perspectives for using classical neural network models and methods of counteracting attacks on network resources of information systems. *News of the national academy of sciences of the republic of Kazakhstan. Series of geology and technical sciences*. 2017. Vol. 5, Number 425. Pp. 202-212.

3. Gamayunov D. Falsifiability of network security research: The good, the bad, and the ugly. In *Proceedings of the 1st ACM SIGPLAN Workshop on Reproducible Research Methodologies and New Publication Models in Computer Engineering, TRUST'14* (pp. 4:1–4:3). ACM New York, NY, USA, 2014.

4. Hu Z., Gnatyuk S., Koval O., Gnatyuk V., Bondarovets S. Anomaly Detection System in Secure Cloud Computing Environment. *International Journal of Computer Network and Information Security*. 2017. Vol. 9, № 4. Pp. 10-21.

5. Zhengbing H., Tereykovskiy I., Tereykovska L., Pogorelov V. Determination of Structural Parameters o f Multilayer, 2017.

6. Toliupa S., Parkhomenko I. Data protection with intellectual support of organizational and technical and operational management. *Radio Electronics and Telecommunication*. 2016. № 3. Pp. 121-130.

7. Toliupa S., Parkhomenko I. The development of a process planning model of rational modular composition of the information protection systems. *Problems of Telecommunications*. 2016. № 3. Pp. 56–64.

8. Toliupa S., Druzhynin V., Parkhomenko I. Signature and statistical analyzers in the cyber attack detection system. *Scientific and Practical Cyber Security Journal (SPCSJ).* September 2018. № 3(02). Pp. 47-53.

9. Toliupa S., Nakonechnyi V., Uspenskyi O. Signature and statistical analyzers in the cyber attack detection system. *Information technology and security. Ukrainian research papers collection.* 2019. Vol. 7, Issue 1 (12). Pp. 69-79.

10. Mohammad Sazzadul Hoque, Md. Abdul Mukit, Md., Abu Naser Bikas. An implementation of intrusion detection system using genetic algorithm. *International Journal of Network Security & Its Applications (IJNSA)*. Sylhet, 2012. Vol. 4, No. 2. Pp. 109-120.

11. Analysis and Evaluation of Network-Based Intrusion Detectionand Prevention System in an Enterprise Network Using Snort Freeware / O. B. Lawal [et al.]. *African Journal of Computing & ICT*. Ibadan, 2013. Vol. 6, No. 2. Pp. 169-18.

References

1. Tereikovskiy, I, Toliupa, S., Parkhomenko, I., Tereikovska, L. (2018). Markov Model of Normal Conduct Template of Computer Systems Network Objects. *14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering TCSET-2018* (pp. 498–501.

2. Aitchanov, B., Korchenko, A., Tereykovskiy, I., Bapiyev, I. (2017) Perspectives for using classical neural network models and methods of counteracting attacks on network resources of information systems. *News of the national academy of sciences of the republic of Kazakhstan. Series of geology and technical sciences, 5*(425), pp. 202-212.

3. Gamayunov, D. (2014). Falsifiability of network security research: The good, the bad, and the ugly. In Proceedings of the *1st ACM SIGPLAN Workshop on Reproducible Research Methodologies and New Publication Models in Computer Engineering, TRUST'14* (pp. 4:1–4:3). ACM New York.

4. Hu, Z., Gnatyuk, S., Koval, O., Gnatyuk, V., Bondarovets, S. (2017). Anomaly Detection System in Secure Cloud Computing Environment. *International Journal of Computer Network and Information Security*, *9*(4), pp. 10-21.

5. Zhengbing, H., Tereykovskiy, I., Tereykovska, L., Pogorelov, V. (2017). *Determination o f Structural Parameters of Multilayer.*

6. Toliupa, S., Parkhomenko, I. (2016). Data protection with intellectual support of organizational and technical and operational management. *Radio Electronics and Telecommunication*, (3), pp. 121-130.

7. Toliupa, S., Parkhomenko, I. (2016). The development of a process planning model of rational modular composition of the information protection systems. *Problems of Telecommunications*, (3), pp. 56–64.

8. Toliupa, S., Druzhynin, V., Parkhomenko I. (2018). Signature and statistical analyzers in the cyber attack detection system. *Scientific and Practical Cyber Security Journal (SPCSJ),* (3(02)), pp. 47-53.

9. Toliupa, S., Nakonechnyi, V., Uspenskyi, O. (2019). Signature and statistical analyzers in the cyber attack detection system. Information technology and security. *Ukrainian research papers collection, 7*(1(12)), pp. 69-79.

10. Mohammad Sazzadul Hoque, Md. Abdul Mukit, Md., Abu Naser Bikas (2012). An implementation of intrusion detection system using genetic algorithm. *International Journal of Network Security & Its Applications (IJNSA), 4*(2), pp. 109-120.

11. Lawal, O. B. et al. (2013). Analysis and Evaluation of Network-Based Intrusion Detectionand Prevention System in an Enterprise Network Using Snort Freeware. *African Journal of Computing & ICT*, 6(2), pp. 169-18.

UDC 004.056.5:004.7(083.94)

Serhii Toliupa, Ivan Parkhomenko, Liudmyla Tereikovska, Volodymyr Kvasnikov

CONSTRUCTION OF CYBER ATTACK DETECTION SYSTEMS WITH THE HIDDEN MARKOV MODEL

One of the most promising ways to improve the quality of data analysis is to use the anomaly detection method in network cyberattack detection systems. In this method, the work of data analyzers is based on the assumption that a sign of a cyberattack is a certain deviation of the controlled parameters of a computer system (CS) from the parameters that characterize the normal functioning of the network. The values of the monitored parameters under normal operating conditions are called patterns of normal behavior. That is why the choice or formation of such a template that would adequately reproduce the functional portrait of the network object of the CS and allow to determine the anomalous behavior of this object with a given accuracy is a very important task.

Massive cyberattacks initiate the creation of special technical solutions, means and systems of counteraction. To detect network intrusions, modern methods, models, tools and complex technical solutions for intrusion detection and prevention systems are used, which can remain effective when new or modified types of cyber threats appear. But in fact, when new threats and anomalies appear, generated by attacking actions with unidentified or vaguely defined properties, these tools do not always remain effective and require long time resources for their appropriate adaptation. Therefore, intrusion detection systems (IDSs) must be continually researched and improved to ensure their effective continuity.

The issue of security and protection of information systems (IS) has been studied by domestic and foreign researchers, some of them are T. Ptaceka, O. Camp, P. Albers, I. Tereykovsky, A. Korchenko, V. Buryachok, V. Dudikevich and other. Undoubtedly, the relevant direction in the field of information security, which has intensive development, is the direction of detecting cyberattacks and protection against interference in the work of information systems by unauthorized parties. In addition, it should be noted that attacks on information systems are becoming more frequent, methods of their implementation are becoming more sophisticated, and the scale is increasingly global.

If we talk about intrusion and attack detection systems, then of course they have numerous flaws in terms of security solutions. In order to increase the efficiency of these systems, not only in terms of detecting harmful effects on protected objects of information systems infrastructure, but also it is necessary to take into account the factors of the daily operation of these tools. In addition, an important issue is economic efficiency, given the optimization of information resources of the owner of the protection system. One of the most effective methods for detecting intrusions and attacks is the method based on the signature approach. As for signature-based methods for identifying attacks, they form a set of rules or a formal model for describing attacks. Regarding the formal model, in this case it can use a character string and a semantic expression in a special language, etc. The main mechanism of the signature method is the use of signatures (a specialized database of certain patterns) of attacks. These signatures are used to search for actions that show signs of an "attack".

As a result of the research, the possibility of forming patterns of normal behavior of network objects of computer systems based on a homogeneous Markov chain with successive transitions has been substantiated. The structure of the model has been developed, a corresponding mathematical instruments have been formed.

An important area of further research is the development of model dependencies of the overall change in parameters that determine the safety of the CS. In addition, the development of patterns of normal behavior of the CS is inappropriate without substantiating the nomenclature and method for assessing the controlled parameters. So, another important area of further research should be the development of a system for selecting and assessing the safety parameters of the CS.

Keywords: *network cyberattack; computer system; information security; intrusion detection system; Markov chain; Markov process; Markov model; characteristic of security.*

Fig.: 4. References: 11.

Толюпа Сергій Васильович – доктор технічних наук, доцент, професор кафедри кібербезпеки та захисту інформації, Київський національний університет імені Тараса Шевченка (вул. Володимирська, 60, м. Київ, 01033, Україна).
Toliupa Serhii – Doctor of Technical Sciences, Professor, Professor of Cybersecurity and Information Protection department, Taras Shevchenko National University of Kyiv (60 Volodymyrska Str., 01033 Kyiv, Ukraine).
E-mail: tolupa@i.ua
ORCID: https://orcid.org/0000-0002-1919-9174
SCOPUS authorID: 57201779357

Пархоменко Іван Іванович – кандидат технічних наук, доцент, доцент кафедри кібербезпеки та захисту інформації, Київський національний університет імені Тараса Шевченка (вул. Володимирська, 60, м. Київ, 01033, Україна).
Parkhomenko Ivan – PhD in Technical Science, Associate Professor, Associate Professor of Cybersecurity and Information Protection department, Taras Shevchenko National University of Kyiv (60 Volodymyrska Str., 01033 Kyiv, Ukraine).
E-mail: parkh08@ukr.net , parkh08@gmail.com
ORCID: https://orcid.org/0000-0001-6889-9284
Web of Science ResearcherID: R-1283-2017
SCOPUS authorID: 57194039632

Терейковська Людмила Олексіївна – кандидат технічних наук, доцент, доцент кафедри інформаційних технологій проектування та прикладної математики, Київський національний університет будівництва і архітектури (Повітрофлотський проспект, 31, м. Київ, 03037, Україна).
Tereikovska Liudmyla – PhD in Technical Science, Associate Professor, Associate Professor of the department of information technologies of design and applied mathematics, Kyiv National University of Construction and Architecture (31 Vozdukhoflotsky Av., 03037 Kyiv, Ukraine).
E-mail: tereikovskal@ukr.net
ORCID: https://orcid.org/0000-0002-8830-0790
Web of Science ResearcherID: V-7948-2018
SCOPUS authorID: 57198815503

Квасніков Володимир Павлович – доктор технічних наук, професор, завідувач кафедри Комп'ютеризованих електротехнічних систем та технологій, Національний авіаційний університет (просп. Любомира Гузара, 1, м. Київ, 03058, Україна).
Kvasnikov Volodymyr – Doctor of Technical Sciences, Professor, Head of department Computerized electrical systems and technologies, National Aviation University (1 Lubomir Guzara Av., 03058 Kyiv, Ukraine).
E-mail: kvp@nau.edu.ua
ORCID: https://orcid.org/0000-0003-3888-772X або https://orcid.org/0000-0002-6525-9721
SCOPUS authorID: 56871189000